Evaluating the Impact of Co-Design

Building a Co-Design Evaluation Framework for Service Development Projects Starting from Buddyschool Case Study.

Alessia Menegolli

Master's Thesis, 30ECTS

Collaborative and Industrial Design
School of Arts, Design and Architecture

Aalto University

2020

Blank Page

Evaluating the Impact of Co-Design

Building a Co-Design Evaluation Framework for Service Development Projects Starting from Buddyschool Case Study.

Alessia Menegolli

Master's Thesis, 30ECTS

Collaborative and Industrial Design
School of Arts, Design and Architecture

Aalto University

2020

Blank Page

Acknowledgements

I would like to thank:

My advisor and supervisor Núria Solsona Caba, for always helping me in finding the way forward.

Migrant Youth Helsinki staff, for allowing me to utilize their precious time and resources.

My friends and colleagues at Aalto University, for the precious support received throughout all the process.

My family, for believing in me and always encouraging me to do what I love and be who I want to be.

Michele, for always supporting me, even from far away.

Abstract

While metrics, indicators and parameters are certainly not missing from current design practices, little research has been conducted to investigate the opportunities for co-design evaluation. Previous literature (amongst others: Kujala, 2003; Steen et al., 2011; Suominen & Pöyry-Lassila, 2013) has focused on identifying the beneficial effects caused by undertaking a co-design approach in a service development project, but at the moment there are no comprehensive studies investigating the impact that such benefits create in the service delivery and performance. Nevertheless, the lack of further research on the topic is becoming more widely acknowledged and several authors are highlighting the need for addressing co-design evaluation opportunities (Steen et al., 2011; Hoyer et al., 2010; Ostrom et al., 2010; Yin et al., 2011). Consequently, the focus of this master´s thesis is exploring the research gap individuated concerning co-design evaluation. Due to the lack of previous inquiries on the subject, a case study is selected to act as a data collection scenario to generate valuable knowledge on the topic. Therefore, this thesis utilizes qualitative research methods in an empirical setting in order to produce further learnings on the research gap identified.

The case study chosen is Buddyschool, a peer-to-peer learning support service co-developed by Migrant Youth Helsinki and relevant actors in the education environment. Buddyschool aims at helping children who are struggling in coping with school by pairing young pupils with older student tutors. At the moment of the research, Buddyschool is implemented in 38 comprehensive schools in the Helsinki region, offering a perfect occasion for researching both the co-design process and the service results, including potential impacts that the co-development process had on the final outcome.

By crystallizing the findings gathered from the background research and the case study´s fieldwork, this master´s thesis proposes a framework for co-design evaluation in service development projects. The framework organizes several metrics recognized throughout the research and clusters them into meaningful categories. Moreover, this thesis argues for the possibility of evaluating co-design both during the development phase and after the implementation of the service. Building on the works of Voss (1992) and Foglieni, Villari & Maffei (2018), the current design proposal can be a tool for achieving consequential learnings through outcome-focused metrics, as well as making adjustments *in itinere* thanks to process-focused metrics. In fact, during the development phase the focus of the framework´s indicators is assessing the collaborative performance and the co-design process, whereas the metrics featured in the implementation phase aim at evaluating the effects of the co-development process on the service performance.

Finally, the value that the framework adds to the practice is compiling a list of relevant indicators, taking a step towards the realization of a comprehensive co-design evaluation framework. In addition, the current design proposal aims at sparking thoughts and reflections on the subject, as well as fostering interest in further researching it.

Keywords

Co-Design, Co-Design Evaluation,
Service Design

Table of Contents

Blank Page

Chapter 01

Introduction

1.1 Background

"The application of participatory design practices (both at the moment of idea generation and continuing throughout the design process at all key moments of decision) to very largescale problems will change design and may change the world"
(Sanders & Stappers, 2008, p. 9).

Co-design is the subject of increasing interest from companies operating both in the public and private sectors. The people formerly known as "users" are now engaged as partners and collaborators and involving them throughout the whole design process seems to have become a priority in most cases. Co-design is seen as essential in the service design practice as it offers methods and tools to bring different perspectives on the table and combine them to create valuable and successful services, enabling a wide away of benefits for the project (Steen et al., 2011). Previous literature has shed light on this array of effects that a collaborative approach helps to implement, giving businesses and practitioners an understanding of the opportunities, that co-design can actualize. Effect include, for instance, reduction in costs (Hoyer et al., 2010), an increase in customers' loyalty (Alam, 2012), changes in the company's mental models (Suominen & Pöyry-Lassila, 2013), and a more accurate fit with users' needs (Kristensson, Magnusson & Matthing, 2002).
Although the benefits of co-design are the subject of different studies in the field of service design, there are no comprehensive researches investigating the impact that the benefits brought about by the co-design process have on the service delivery and service performance.
Moreover, further research is needed on possibilities to measure such effects and the impact that co-design has on the project. This need for further inquiry is also underlined by the under-researched topic of service design evaluation. In fact, the literature contributes mainly on service evaluation, which focuses on measuring the results of the service, while there is considerably less research on service design evaluation focusing on understanding and assessing the impact of the co-design process on the service performance.
In summary, while previous literature has focused on defining co-design through definitions, characteristics and best practices (Blomkamp, 2018; Sanders, 2002; Sanders & Stappers, 2008; Steen, 2013), as well as underlining the effects and benefits of co-design (Kujala, 2003; Steen et al., 2011; Suominen & Pöyry-Lassila, 2013), further research on the topic of evaluation concerning service design and co-design is needed (Hoyer et al., 2010; Ostrom et al., 2010; Steen et al., 2011).

This thesis argues that evaluating the impact of co-design is needed for three main reasons.
Firstly, showcasing the impacts of co-design is useful in order to advocate for a bigger role of co-design inside organizations by offering proof of the link between a co-design approach and the benefits delivered in the service performance.
Secondly, evaluating the process, and not just the outcomes, can foster reflective activities aimed at improving practices and learning. By employing a continuous evaluation, the learning will not only be applicable on future projects, but also on the current one, allowing the co-design team to make changes while developing the service, maximizing the potential of the service implementation.
Lastly, measuring the co-design process is useful for accountability reasons. Evidence of the process and its effects on the service delivery will help in showing justified choices, based on real and reliable data, supporting the utilization of resources by the organization.
Consequently, the aim of this thesis is exploring the subject of co-design evaluation. This study argues for the possibility to assess the collaborative performance and the co-design process during both the development phase and the implementation phase. The development phase encompasses all the activities that lead to the implementation of the service, such as research,

co-design activities like workshops, interviews and service prototypes, and various iterations. On the other hand, the implementation phase concerns the service after the executive buy-in, when it is brought to life and implemented in the environment it was thought for.
During the development phase the focus of the assessment can be found in evaluating the collaborative performance, offering opportunities for making changes *in itinere*, as well as gathering data to enable future reflections. Similarly, in the execution phase the evaluation can highlight the effects brought about by co-design in the service performance and create a link between co-design benefits and the service delivery.
In order to investigate the opportunities of co-design evaluation, a specific case study is selected to act as a data collection scenario. The case study featured in this research is Buddyschool, a service born from a 4-years-long co-design process undertaken by Migrant Youth Helsinki.
Buddyschool is a peer to peer learning support developed to address the challenges of students who are enrolled in any institution of the Finnish education system. Buddyschool was chosen because of its co-design approach that allowed Migrant Youth Helsinki to include several users and prototypes and workshops. In addition, the service is currently implemented in 38 schools in the Helsinki region, which allowed the research to investigate the effects of the process on the actual service results.
A more in-depth understanding of the case study and its co-design process is offered in chapter *2. Research Setting*.
The research outcome presented in this master´s thesis is the result of the analysis of relevant literature and the Buddyschool case study´s fieldwork. The design proposal is a framework compiling different metrics gathered from both the literature and the field research.
Such framework is proposed as the first step towards a categorization of metrics that could be used in co-design evalutation. The value that this framework adds to previous research and current practice is providing a tool that features an organized synthesis of co-design evaluation criteria, as well as sparking interest in the topic and in potential future research.

1.2 Type of Thesis

In order to frame the research of the wider topic on a more empirical level, a specific case study was selected for the research. Therefore the research consists of both field-based and literature-based research.
In general, the overarching methodology for this master's thesis can be identified in Action Research, which is relevant to this thesis because it stresses reflective practice, as well as closely working with a partner organization in the field, conducting the esearch with people, rather than for or on them (Coghlan & Brannick, 2005, p. 4).
The two layers of this research, the case study and the background research with literature review and expert's interviews, have influenced each other continuously. The fieldwork in the case study was helpful to determine the final outcome, and in turn the background research was helpful to inform how a general framework could be applied to the specific case study.
Further discussion on the Buddyschool case study and Action Research as a methodology will be presented in chapter *3. Methodology and Methods*.

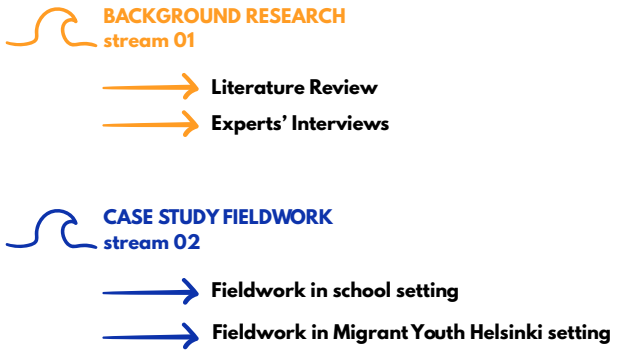


Figure 1. The two research streams developed in this master´s thesis.

1.3 Research Objectives

The personal motivation in researching this topic was fostered especially during discussions in academic settings. During Autumn 2019, a group of students worked closely with the City of Helsinki on the topic "evaluating the impact of design". The outcome of one of the groups in the course "Strategic Co-design" from Aalto University was a map showcasing the several design projects in the City divided by category. This project sparked the interest of further researching the topic of evalutation and design, but a different focus and lens were chosen to approach the topic, deciding to focus on exploring co-desing evaluation through Buddyschool case study.

In general, this master´s thesis can be considered of relevance in the service design and co-design practice because it explores an under-researched area such as co-design evaluation. Indeed, the interest for this topic was also sparked due to the excited new horizons offered by the not-so-populated research and literature. Given the lack of consistent previous references, this research choses to use a case study which offers an empirical approach to the subject. This offered the possibility to collect data on the process, as well as on results of the service implementation. In addition to researching the case study, a general inquiry on the state of evaluation in the service design and co-design practice was carried out though literature review and experts' interview.

Thus, the research objectives are:

1. Explore the current discourses of evaluation in co-design applied to service design. What is the state of the art? Is there a real need for evaluating the impacts of co-design? What are the motivations? Are there evaluation methods or metrics already in use?
2. Research the case study of Buddyschool to obtain empirical data on the topic based on a real-life setting. Are the general discourses on evaluating the impact of co-design applicable to Buddyschool? How can Buddyschool be helpful for researching this topic?
3. Research the possibility of designing an evaluation framework applied to Buddyschool. Then, discuss if and how this could be applied to other services.

1.4 Research Questions

In order to conduct this research, the thesis utilizes qualitative research methods to investigate on two layers. In addition to the case study already mentioned, an overview of discourses concerning the topic was gathered through a series of four experts' interview and a review of relevant literature.

Therefore, the research questions are:

1. What is the state of the art on evaluation in co-design?
    - 1.1 What are the current discourses on evaluating the impact of co-design?
    - 1.2 Which methods and metrics are used to evaluate and measure co-design in a service design project?
2. In Buddyschool, how can we evaluate the impact of co-design?
    - 2.1. Which methods can be used to measure the outcomes and the co-design process of Buddyschool?
3. Can we scale the findings from Buddyschool to other services?
    - 3.1. What are the limitations?

1.5 Research Scope

The scope of the research is narrowed down to researching the topic of evaluating the impact of co-design in a specific case study, Buddyschool. Other studies (see, e.g., (Björklund et al., 2018; Foglieni, Villari & Maffei, 2018; Foglieni & Holmlid, 2017)) take a more general approach and discuss evaluating the impact of service design or design thinking. While the research presented in this master's thesis deals specifically with collaborative design, regarded mainly as the

involvement of users and relevant stakeholders in the service development process, other studies will be taken into consideration in the literature review and through the expert's interviews, as they serve as context and inquiry into established practices.

This master´s thesis considers co-design as an integral part of the service design development process. In this sense, the various research streams encompassed in this master´s thesis introduce multiple references to service design works and literature. Mainly due to the lack of research delving into the area of co-design evaluation, service design literature is used in an explorative way, becoming a baseline from which to develop proposals more specific to co-design.

In other words, service design is used as a lens for exploring the evaluation discourses within service design practice and research. Successively, the findings related to service design evaluation areas are transformed and applied to the co-design development process. In this way it is possible for the research to contextualize the findings in a wider discourse, while still researching on the specific topic of evaluating the impact of co-design. The most evident example of the process of revising and reshaping elements and theories of service design evaluation to find and highlight co-design evaluation opportunities is the re-adaptation of the work by Foglieni, Villari & Maffei (2018) dealing with service and service design evaluation to create the fundamental structure of the design proposal featured in this thesis, an evaluation framework for co-design in service development projects. Further discussion on this matter is presented in chapters *5. Analysis and Findings* and *6. Research Conslusions and Proposals*. In addition, also Migrant Youth Helsinki staff recognizes co-design in Buddyschool as part of a broader service design approach.

Figure 2. Service and Co-design evaluation as used in this master´s thesis.

1.6 Thesis Structure

In this chapter, *1. Introduction*, an overview on the topic is presented, as well as the research questions and objectives. The following chapter, *2. Research Setting*, introduces the context of the case study by presenting the Buddyschool service and its co-design process. The third chapter, *3. Methodology and Methods*, presents an overview of the data collection and analysis methods. The fourth chapter, *4. Literature Review*, explores relevant literature around co-design evaluation. An introduction to co-design is given, as well as discussing the effects of co-design in a service design project, evaluation practices and metrics. Next, chapter *5. Analysis and Findings* presents the findings of the research and sets the stage for a discussion which is further elaborated in the following chapter, *6. Research Conslusions and Proposal.* By tying together the findings from the different research streams, the sixth chapter introduces the design proposal, a co-design evaluation framework, which is further discussed in all its parts and applied to the Buddyschool case study. In conclusion, the last chapter, *7. Discussion*, summarizes the research by suggesting future research opportunities, limitations of the current research and by sharing personal reflections on the research journey.

Chapter 02

Research Setting

2.1 Migrant Youth Helsinki and Buddyschool

Migrant Youth Helsinki is a project established in 2016 to address the mission of reducing marginalization and promoting equal opportunities for all youth living in Finland. During the year when the project started, data reported that 14.3% of Helsinki population was composed by citizens with a migrant background. Nevertheless, this number is estimated to grow and reach 23% of Helsinki population by 2030 (Hiekkavuo, 2017).
Despite the high number of people with a migrant background choosing Helsinki as their home, 25% of migrant youth is unemployed and not attending any education institute nor carrying out any non-military or military service.
In contrast, Finnish speaking youth in the same situation is only 4% (Markkanen, 2017).
Migrant Youth Helsinki project was born to address this issue in order to prevent the discrepancy to become permanently embedded in the Finnish society, resulting in a disengagement with education by a whole generation. Migrant Youth Helsinki decided to take a design- and experiment-led approach on solving this challenge and prototyped several services thought *ad-hoc* for bringing benefits to migrant youth.
The four-year-long project was funded by Me Foundation and will come to an end in 2020. The project collaborated closely with a design agency based in Helsinki, Solita, who offered expertise on the collaborative design process.
One of the service ideas born from the co-design process undertaken by Migrant Youth Helsinki project is Buddyschool, a peer to peer learning support developed to address the challenges of students in Finnish education institutions. The model is thought to be scalable from basic education to upper-secondary education (both general and vocational curriculum), but is now being implemented in 38 basic education institutes in Helsinki.
The idea behind Buddyschool is to give the opportunity to older students to mentor and teach younger pupils, providing a chance to improve the learning results of those having difficulties coping with school. Peer learning is not only helpful to boost the academic skills and record of both the younger and the older students taking part in Buddyschool, but also offers an exclusive chance to learn and practice "soft skills" such as self-esteem, sense of responsibility, interaction and communication skills. This translates into a positive impact in the school environment and atmosphere, which in turn have an impact into the positive learning experience of the students, supporting students to continue with their educational path toward an upper-secondary institute.
Buddyschool was chosen as a case study because of its co-design process which faciliated engagement with relevant stakeholders and users. In addition, Buddyschool is now being implemented in schools, meaning that the field research could be able to focus on studying the results of the service and the real-life implementation in the school settings.
Moreover, the project is currently coming to an end, offering an opportunity for creating and maybe implementing practical suggestions on evaluating the impact of co-design in the specific setting of Buddyschool.
The next section describes more in depth the process that generated Buddyschool as a service and the co-design elements in the process, highlighting why Buddyshcool can be considered the result of a co-design effort.

2.2 From idea to service: an overview of Buddyschool development process

Before describing the process that led to Buddyschool´s implementation, it is critical to underline why Buddyschool could be considered as a co-designed service.

The NSW Council of Social Service (2017) argues that the four pillars of co-design can be found in a clear understanding of the problem, a genuine research for solutions that are not

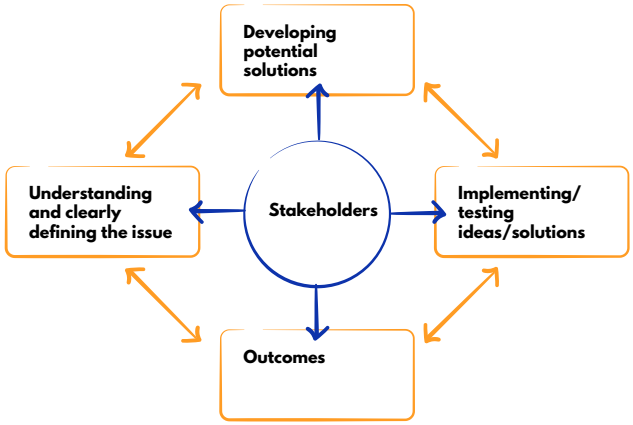


Figure 3. Co-design principles according to NSW Council of Social Service (2017).

pre-conceived, a focus on creating outcomes that serve for the better and a continuous iterative accent. These four elements have one common denominator: the involvement of stakeholders in the process.

All these elements can be found in Buddyschool. First and foremost, Buddyschool is the result of a strong co-development process amongst different stakeholders and users, adding to the project that distinct common denominator found in involving different actors. From the very beginning of the project, the civil servants operating in Migrant Youth Helsinki partnered with a service designer and created a fostering environment for collaboration and co-design, welcoming key stakeholders and users to join the process and to provide meaningful input through their participation. Secondly, as a consequence of the involvement of teachers and actors in the educational environment, the problem statement was discussed in depth and the general brief provided by City of Helsinki was readapted to serve better the scope of the project. In fact, if the initial design brief addressed the possibility of helping the youth to "cope better", the whole Migrant Youth Project´s aim is to clearly define such issue and divide it into actionable areas, such as the education sector.

Thirdly, one of the main focuses throughout all the process was ensuring that the real needs of the users would be taken into account. The initial phase focused on investigating the genuine needs and setting of the users, without promoting any pre-conception or pre-made solution to impose on the system, but on the other hand trying to tailor a service that would address the observed challenges.

Furthermore, the willingness to adopt a genuine research approach and the involvement of stakeholders and users that are living the school environment in their everyday lives highlighted the necessity to create a service that would change the dynamics for the better, focusing on creating solutions.

Lastly, Buddyschool process was strongly depicted by an iterative accent. In fact, iterations were carried out not only during the research and design phase thanks to service prototypes, co-design and validation workshops and interviews, but iterative practices are still implemented also after the service implementation thanks to regular check-ins with the teachers in charge of the program in the various schools.

Therefore, it can be argued that Buddyschool is a co-design service because of its iterative process, which involved users and stakeholders, promoted empathic understanding of the situation and fostered joint activities and discussions in co-design settings (such as workshops) where the participants could work together to generate ideas and solutions. Moreover, the real needs of the users were investigated, discarding ready-made solution imposed on the school setting.

After exploring the reasons why Buddyschool can be considered a co-designed service, the following section provides a brief overview on the main steps of Buddyschool development process. In addition, the process here presented further underlines the co-design principles described above.

Buddyschool´s process can be divided into two phases: a design phase represented by exploring the brief, brainstorming ideas and service prototyping, and an execution phase, characterized by goal setting, the actual service

implementation, and the scaling of the project. The division line between the two phases is represented by the executive buy-in. The representation of these Buddyschool´s phases is a result of the interviews conducted with Migrant Youth Helsinki staff.

Design Phase.

1.Identify the needs and problem.
The project builds on previous research based on evidence that education enhances children's future opportunities. At the beginning, the brief was broad: the request was exploring ways on how to create a positive impact on migrant youth and offer them better integration in Finland and support their future lives in this country.
To scale down the brief and explore the design space, Migrant Youth Helsinki held several workshops aimed at investigating different areas and themes. The topics explored were education, working life, and social landscape. Every topic had its own set of workshops and working groups, which comprehended experts in the field. For instance, in the group aimed at exploring the education setting, principals, teachers, psychologists, and youth workers were present, as well as the project manager of Migrant Youth Helsinki and the lead service designer of the project. In general, the aim of these workshops was to define a scope and start discussing about opportunities and possibilities for design interventions.

2.Idea generation
Once the opportunities' space was defined, a new series of workshop was held with the goal of generating ideas. In these workshops the same actors participating the "scope-definition" workshops were present.
The co-design groups generated several ideas relevant to their own area of expertise through the facilitation of the project manager and the service designer.

3.Validate problem-solution fit.
Feedback was then asked to a group of "young-designers". The young migrants were asked to give their opinion on the
ideas and share their perspectives, as well as investigating what their friends would think about the concepts. At the same time, also the "adult co-designers" tested the ideas with their own peers, colleagues, and customers. For instance, in the education group, youth workers explored the concepts emerged during the workshops with other youth workers but also with the children they were working with.

4.Try in the real world.
After gathering the feedback on the specific ideas collected in the workshops, the next step is selecting the most promising ones based on the feedback and prototype them. For instance, in the education group, the idea that let to Buddyschool was chosen as the one to be tried out in a real-life setting. The initial concept for what is now Buddyschool was quite simple and can be summed up with the sentence "help for children to do homework, no adults allowed".
Certainly, the current service delivery is different in many ways from this initial framing, and service prototyping played an important role in adjusting the first idea into a full-working service.
Buddyschool was tested in several schools, prototyping different opportunities for implementation. Feedback from the teachers and the children was gathered to assess the various settings.
The process was iterative, meaning that the real-world experiments informed further changes in the idea, which in turn required to collect more feedback and prototype again. In general, the experiments were deemed as a crucial part of the process since they provided means to foresee implementation and scaling opportunities or lack thereof in the Helsinki region. In fact, as Buddyschool advanced with service prototyping, more schools and new teachers were involved, expanding the original co-design setting to include more and more people.

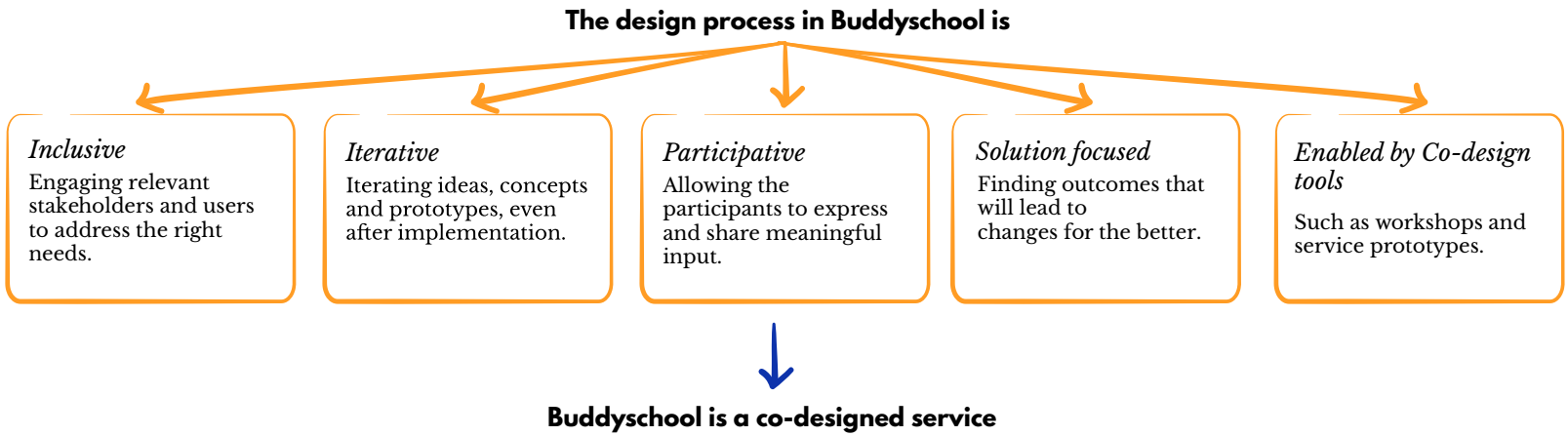


Figure 4. Buddyschool and co-deisgn. Principles adapted from NSW Council of Social Service (2017).

Buy-in at the executive level.

After the iterative process refined the concept in a successful way, the next step is represented by the executive level's buy in. This action represents the point of division between the design phase and the execution phase, which starts with setting up the frame for implementation.

Execution Phase.

5.Recruitment of a product owner.
After obtaining the buy in, a product owner was recruited. This was essential to provide a leading figure in Buddyschool who would collaborate closely with the project manager.

6.KPIs and goal setting.
The key for a successful goal setting according to Migrant Youth Helsinki staff is to be ambitious. Setting ambitious goals was an important factor in Buddyschool´s process because it allowed the product owner to define smaller task in a more accurate way and understand what to prioritize. In addition to that, it highlights the seriousness of the staff and their hard-working attitude.

7.Organization-solution fit.
Alongside the goal setting, testing the fit between the solution and the organization was a big part of the execution level phase. This is a continuous task that needs to be carried out steadily to adapt the service to the different needs of the several schools. If in the design phase the focus was to identify whether the idea and product fit was appropriate, in this phase the goal is to design for the fit of the organization. Feedback was gathered in the form of anecdotes and stories from teachers and questionnaires. In this sense, the value that co-design brought to the design phase is allowing the co-designers to feel part of the process and articulate their needs into ideas, whereas the value that co-design adds in the execution phase can be seen as an "everyday co-design", through which teachers in charge of the service delivery can make small adaptations to Buddyschool's features to fit it in their own school's environment.

8.Scaling.
The scaling is accomplished on two levels. Firstly, by scaling in the number of schools involved and then by starting to scale the size of

Buddyschool service within each school.
Buddyschool is part of the national curriculum and therefore it's a recommended program that each school can decide to join. When a school decides to implement Buddyschool in their own premises, the product owner offers initial support during the first phases. According to Migrant Youth Helsinki staff, the product owner holds training sessions to introduce the teachers to the Buddyschool concept and ways of working. In these trainings a brief overview of the process is given to the teachers.
After providing initial and customized support, the program is left running in the hands of responsible teachers in the school. It is highlighted how the community of teachers formed partially thanks to the co-design process helps in providing further training and sharing best practices. In a way, the peer-to-peer learning characterizing Buddyschool approach towards children is translated to the teachers as well.
As mentioned for the design phase, the execution phase is also identified by strong iterations accents. In this sense, feedback gathered helped in further developing Buddyschool even after the service implementation to make sure the program would fit into the school environment and operations.

Overall, it can be argued that the development process of Buddyschool is strongly characterized by a co-design factor that is translated in both the design and the execution phase.
In addition to that, Migrant Youth Helsinki staff demonstrated competences and capabilities in service design, design thinking and co-design, which were utilized throughout all the process.
The motivations to select Buddyschool as a case study for this master's thesis are the strong presence of co-design accents during the process, the current stage of the project, and the fact that throughout the process there were indeed moments dedicated to evaluation, but none of them focused on evaluating co-design as such.
In this sense, Buddyschool offers a good chance for this research to investigate further the topic of co-design evaluation, as it offers empirical opportunities to articulate specific metrics tailored to Buddyschool and validate the findings of the background research.

ACTORS IN THE BUDDYSCHOOL DEVELOPMENT PROCESS

✳ **Interacted with during the thesis process.**

MIGRANT YOUTH HELSINKI STAFF

Project Manager of Migrant Youth Helsinki ✳

Present throughout the whole process.

Product Owner of Buddyschool ✳

Present from the execution phase.

Service Designer ✳

Present mostly in the design phase. Occasional presence in the execution phase.

FROM SCHOOLS/ EDCUATION SECTOR

Education experts (other than teachers)

They are principals, psychologists, and youth workers. They were present in the problem definition workshops and in the idea generation workshop. They tested the ideas with their colleagues and customers.

Teachers ✳

They were present in the problem definition workshops and in the idea generation workshops. Interviewed and asked feedback to throughout all the process. Participated in the service experiments. The original pool of teachers comprehended 17 participants to the first workshop, but the numbers increased quickly. In the fourth workshop they already had more than 50 teachers present.

Young co-designers

They gave feedback and tested the ideas amongst their peers in the design phase. They expressed their opinions on a variety of services concerning the different thematics explored by the "adult co-designers", not only on the education sector. Young co-designers are 10 young migrants.

Children in schools ✳

They participated in the service experiments and gave feedback on Buddyschool. After the service implementation, the teachers are in charge of collecting feedback from them and report it back to the product owner.

OTHERS

Steerring group

Approved 5 projects for further development, including Buddyschool. The steering group is chaired by the executive director for culture and leisure for the City of Helsinki and includes members from several city departments, members from We Foundation, and two young migrants from the original 10 young co-designers.

Figure 5. Roles in Buddyschool's process.

2.3 Research Timing and Context

The research preparations started during summer 2019, when Migrant Youth Helsinki's project manager was first contacted in order to test the ground for a possible collaboration.
During Autumn and Winter 2019, several meetings occured in order to explore common interests and directions that the thesis could have. Originally, the main motivation for contacting Migrant Youth Helsinki was working with migrant youth, but given the timeframe of the project, the original idea of exploring co-design practices with migrants was not suited for being the thesis topic.
On the other hand, the opportunity of evaluating the impacts of co-design arose due to the final stage of the project. For the sake of this master's thesis, only one service from Migrant Youth Helsinki's portfolio is taken as focus.
The actual work with the case study started in January. After obtaining a research permit from City of Helsinki and arranging the visits with the teachers, the field research started in February and was carried out in one nine-year compulsory basic comprehensive institute in the Helsinki region, which will be kept anonymous.
Further research was conducted with the staff in the form of interviews. Due to exceptional circumstances caused by the spread of COVID-19, schools closed from mid-March. This translated into a switch to remote working for the final months of the thesis work.
Collaboration with Migrant Youth Helsinki was based on the mutual interest of further exploring and researching Buddyschool program. Migrant Youth Helsinki team was supportive of the research in the initial phase, but after the switch to remote work the communications decreased due to the urgency of dedicating their resources to more current matters.
In general, Migrant Youth Helsinki team offered their support also in management work, helping in arranging school visits and obtaining the research permit. This support was crucial, and the positive atmosphere of the team made the researcher feel welcomed from the very beginning. Nevertheless, the interest of this master's thesis is to use Buddyschool as a case study to further explore the topic of co-design evaluation, and therefore it does not primarily focus on studying the service and its own characteristics as such.
Another key collaboration that allowed the research to advance was the one carried out with a bachelor design student, Anna Vienamo. She assisted in management work by translating needed documents in Finnish. Moreover, she also took part in the field research carried out in the schools, interviewing the children in Finnish, and she took actively part in the data analysis, further discussing thoughts and opinions. Anna's input was crucial to contextualize the research and bring to the table a fresh view on the fieldwork and insights.
Further discussion on the research settings can be found in the next chapter, *3. Methodology and Methods*.

BUDDYSCHOOL: STEPS OF THE DEVELOPMENT PROCESS (design and execution phase) AND STAKEHOLDERS INVOLVEMENT

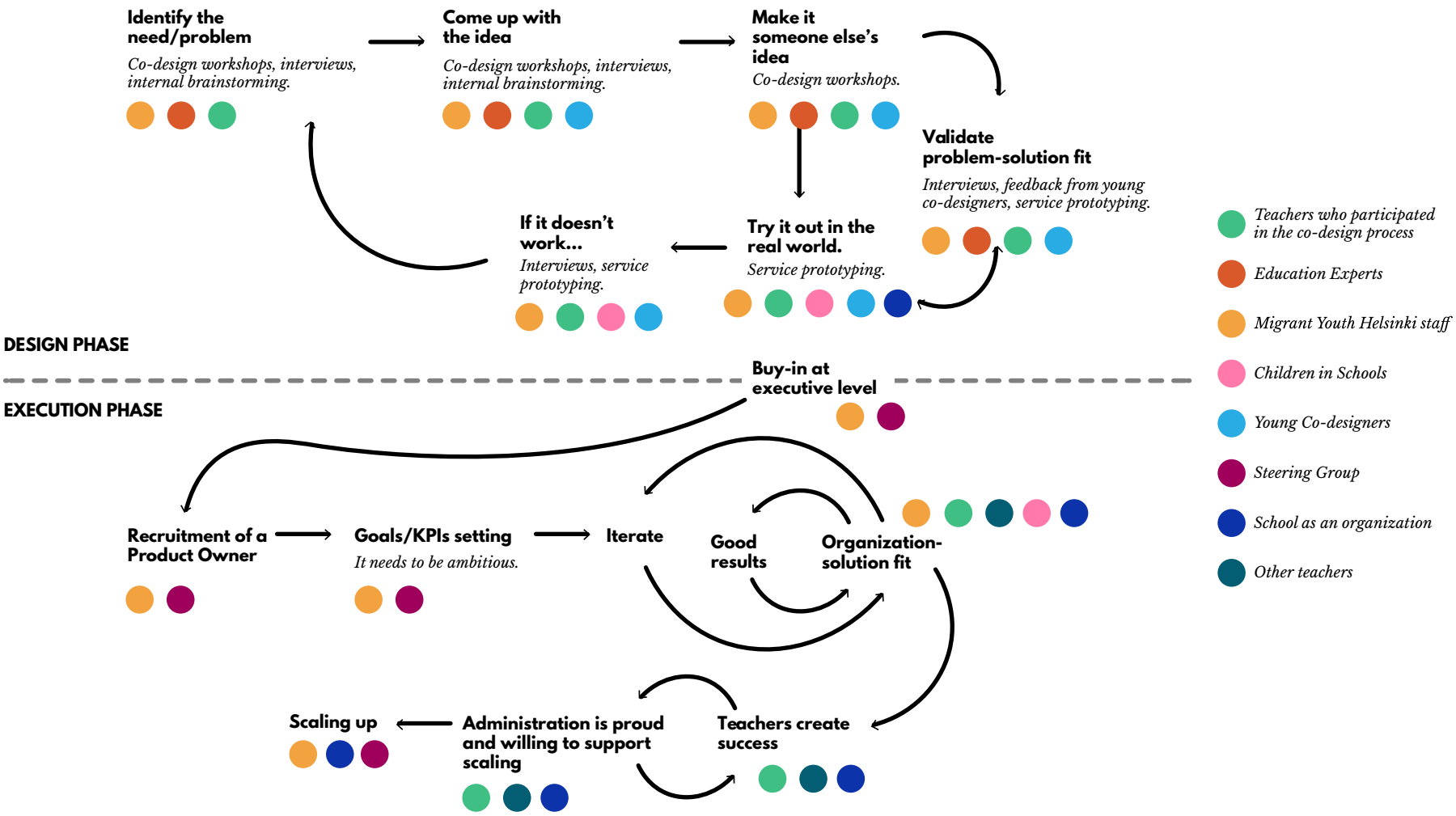


Figure 6. Steps of the process in the Buddyschool development process and mapping the involvement of the key actors. The steps of the process are gathered through the interview with Migrant Youth Helsinki Service Designer.

Chapter 03

Methodology & Methods

3.1 Action Research

This research investigates co-design evaluation through the empirical lenses of the case study.
Since research on this topic is still at the early stage, no substantial prior knowledge has already been generated by academia or practitioners. Therefore, this master´s thesis focuses on action (empirical research), but also on reflection (experimental knowledge production), and has a collaborative approach to reseach. Thus, Action Research is chosen as the overarching methodology.
Coghlan & Brannick (2005, p. xii) define Action Research as an "approach to research which aims at both taking action and creating knowledge or theory about that action. The outcomes are both an action and a research outcome, unlike traditional research approaches which aim at creating knowledge only. . . . The second dimension of action research is that it is collaborative, in that the members of the system which is being studied participate actively in the cyclical process"(Coghlan & Brannick, 2005, p. xii).
A focus on action, reflexivity, collaboration and learnings achieved through the process is also highlighted by Practice Led Research, also addressed by Frayling (1993) as Research through Design (as opposed to Research into Design, and Research for design).
This master's thesis can be contextualized as a Research Through Design in the sense that it utilizes design research methods in order to investigate research questions that are applicable to a real-life setting, but still represent and aim to research the broader current discourse.
Furthermore, Research Through Design and Action Research share fundamental characteristics that result in them being similar (Papas et al., 2012), thus they are both mentioned in this section.
According to Ritchie (2003), qualitative research methods are "of particular value where behaviors and interactions (whether acted, spoken or written) need to be understood in 'real world' contexts" (Ritchie, 2003, p. 34). Therefore, the fieldwork presented in this master's thesis is qualitative because it is related to a setting tied to a real-life environment, the classroom. Moreover, qualitative research allowed to capture more nuanced views and data.
As mentioned before, Action Research´s focal point is collaboration, as it focuses on research with people, rather than for or on them (Coghlan & Brannick, 2005, p. 4). The collaborative approach to inquiry is reflected in this master's thesis by partnering with relevant stakeholders and users throughout all the research process. Stakeholders include the staff from Migrant Youth Helsinki, with whom the researcher collaborated not only during the field work but also in the early stages of the problem definition, the school staff and children and the people who supported the research, such as the Finnish researcher that helped in carrying out the fieldwork. In this sense, the goal has always been to research with people in order to co-produce valuable knowledge and insights that could be translated into action or action plans.
In addition, reflection was a big part of the research. The reflection was both on the topic investigated but first and foremost a personal reflections on the researcher's own journey in the academic and professional worlds. A more thoughtful and extended consideration on the reflective practice throughout the thesis process will be presented in chapter seven, *7. Discussion*.

Concerning Action Research, the model that this master's thesis refers to is specifically Denscombe's Action Model (Denscombe, 1998). He introduces the four main components of Action Research (action, research, collaboration and reflexivity) in a cyclical process. The five steps of Denscombe's model are represented by professional practice, critical reflection (identification of the problem), research (systematic and rigorous enquiry), strategic planning (translation of findings into an action plan) and action (instigate change). A graphical representation is featured in the work by Costello (2011, p. 12). Denscombe's model is relevant to the approach

of this master's thesis and resemblances can be found in several steps. Firstly, the interest in researching the topic of evaluating the impacts of co-design arises from personal experience and professional practice. Working experience in a service and strategic design agency and Aalto University's courses such as "Designing for Services" and "Strategic Co-Design" informed the decision of taking the topic idea further and research it through this master's thesis. Secondly, the critical reflection step coincides with the literature review and the experts' interviews. After gathering knowledge useful to have a critical stand on the matter, the fieldwork started, represented by "research" in Denscombe's model. After that, the strategic planning is expressed by the analysis of the field work to translate the findings into insights and eventually into an "action plan" or design proposal. Finally, the action is represented by the knowledge input shared with relevant stakeholders. The aim is to instigate actions as a consequence of the research findings, for example a wider reflection by the Migrant Youth Helsinki team to implement an evaluation strategy. In general, the reflective part permeated the whole process.

Though no financial compensation was received, nor the researcher was officially part of the team, a fruitful collaboration was achieved through several meetings and the openness of Migrant Youth Helsinki team to welcome new takes on their projects coming from an external party. Thus, a "research in action rather than about action" (Coughlan & Coghlan, 2002, p. 222) was achieved, as the researcher felt part of the Migrant Youth Helsinki and Buddyschool team to some extent, acting partially on their behalf when conducting the fieldwork and immersing herself in the case study. In addition to this, all the knowledge and findings were always shared with them.

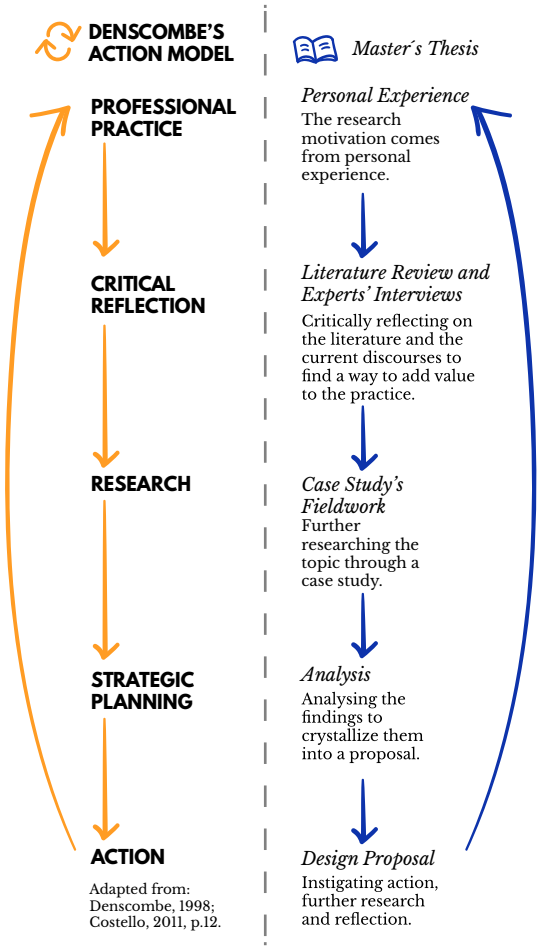


Figure 7. Denscombe´s Action Model as visualized by Costello, (2011) (on the left), mapped against the thesis process (on the right).

3.1.i Barriers to Action Research

It is recognized that Action Research critics argue for the inability to scale the results obtained through action research. Action Research is always context-driven, therefore it can provide valuable knowledge in a specific context, but the possibility of scaling the findings remain questionable, meaning that there is a "split between general theory and local experience" (Gustavsen, 2008, p. 425).
Even if the validity of scaling the results remain unanswered, as well as the applicability of findings to other cases, action research generates actionable and valuable knowledge. In addition to that, the topic of co-design evaluation is still underexplored, thus the reflective character of this methodology will produce valuable explorative knowledge and learnings, regardless of the outcome of the research or its scalability.

3.2 Data Collection Methods

The methods employed for data gathering during the field research are three, observations, field notes and semi-structured interviews. The field research is composed by two streams, one being represented by the case study and the other one being the background research.

Moreover, in the case study stream we can find a further sub-division between inquires carried out with the Migrant Youth Helsinki team and the research in the school environment.
Concerning the school setting, one comprehensive education institute based in Helsinki was selected as the main ground for fieldwork. The school remains anonymous in this research, as all the students and teachers who took part in interviews and observations. The school was selected with the aid of Migrant Youth Helsinki staff by taking into consideration the involvement of the teachers in the project, the quickness of the principal to give approval for the research, and the availability of Buddyschool classes for the researchers to attend to.
During the fieldwork, the researchers attended one Buddyschool class for about one hour at the beginning of March. Buddyschool classes are dedicated times to the Buddyschool program, where children from different ages come together to either be pupils or tutors. During the class a teacher was present the whole time. Ten children from second grade and four children from sixth grade were present. The class consisted in activities such as reading and math exercises.
Observations were used mainly in the school environment, whereas semi-structured interviews were used when talking with experts, Migrant Youth Helsinki staff, and in the schools.
Field notes were used extensively throughout the whole process.
After collecting data on the field, knowledge and insights were produced by reflecting on such data, most of the times in a collaborative manner by confronting ideas with the Finnish researcher that participated in the fieldwork, or with Helsinki Migrant Youth staff, or with the thesis' advisor and peers at Aalto University. In this sense, the knowledge produced is the result of re-elaborating personal views and assumptions by collaborating with several people.

3.2.i Semi-Structured Interviews

Interviews "provide an opportunity for detailed investigation of people's personal perspective, for in-depth understanding of the personal context . . . and for very detailed subject coverage" (Ritchie, 2003, p. 36). Semi-structured interviews are one of the main methods chosen because they allow a flexible approach, even though the structure is prepared beforehand (Legard, Keegan & Ward, 2003, p. 141).
Before conducting the interviews, a topic guide was prepared, highlighting the ideas and topics that were important to address during the session. Nevertheless, it was crucial to engage with the participant and encourage them to talk in an

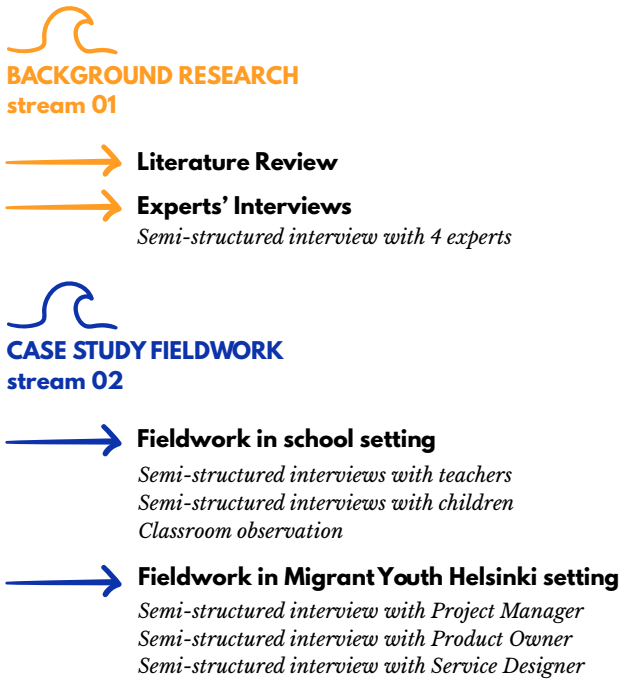


Figure 8. The two research streams developed in this master´s thesis.

open manner about the topic. In this way it was possible to achieve interesting findings that the questions were not aiming for. This is particularly true in the experts' interviews, as the participants were clearly expert practitioners or researchers and held a greater knowledge on the subject than the interviewer. This allowed the conversation to be guided by both parties at the same time, stirring towards interesting and new directions.

All the interviews were structured in a similar way. Firstly, an introduction to the topic and the master's thesis research was given, as well as a brief introduction about the interviewer. Secondly, introductory or "warm up" questions were asked, for instance by asking the participant to introduce herself and her practice. After that, the interview moved on to the main body of inquiry with different aims and questions depending on the participant and the setting. Finally, closing questions were asked, as well as giving the possibility to the participant to add or ask questions about the research or about topics that were not touched upon in the interviews.

The method of semi-structured interview was used in three different settings throughout the research.

1. Semi-structured interviews with experts.

Four interviews were held with experts in the service design field, both from the practice and the academia contexts. The aim of these interviews was gathering knowledge about the current discourses in service design evaluation, in order to have a critical basis and understanding before starting the data collection inherent to the case study. The questions asked were the same during the initial three interviews, with only a few tailored questions concerning articles or researchers of the experts in question. These first three interviews were treated as data generation and collection, whereas the fourth interview was used as a validation interview of the findings previously gathered with the other three sessions.

2. Semi-structured interviews with Migrant Youth Helsinki staff.

Three official interviews were held with Migrant Youth Helsinki staff. In addition to that, there were multiple occasions for meeting with the project owner of Buddyschool and the project manager. The interviews' questions were similar, but they focused on the expertise of the participant. For example, when talking with the service designer, inquiries about the co-design process were highlighted, but when talking with the project owner, more space was given to questions about the Buddyschool implementation and environment.

3. Semi-structured interviews carried out in

schools with children and teachers. The fieldwork featured one school visit, during
which three teachers were briefly interviews, as well as 14 students (10 students from 2nd grade and 4 students from 6th grade). The questions for the teachers focused on finding opportunities for evaluation and inquiry about Buddyschool process and environment in their own school. The questions for children were focused on investigating their experience and thoughts about Buddyschool.

3.2.ii Observations

Observations offer "the opportunity to record and analyze behavior and interactions as they occur" (Ritchie, 2003, p. 35). Observations were used as a main research method during the school visit. In order to minimize the intrusion in the setting, observations were preferred in the initial phase to capture non-verbal communications and the atmosphere of the environment. Because Buddyschool was taking place in Finnish, the English-speaking researcher had the opportunity to focus more on observing the setting, the relationships and the body language, while the Finnish speaking researcher focused on listening to spoken interactions. This allowed to gather both verbal and non-verbal actions.

3.2.iii Field Notes

In order to crystallize observations, field notes were used. In addition to serve as means for capturing observations and interviews' discussions, field notes were also used in general to document thoughts sparked from the research process or from other events or conversations concerning the topic of evaluating the impacts of co-design. Therefore, field notes could be categorized in two typologies: field notes coming from events external to the case study that helped to inform choices and generating knowledge and reflections, and field notes coming directly from the research, crystallizing observations, interviews, thoughts or conversations occurred with relevant stakeholders in the process. In this sense, fieldnotes are "notes made by researchers 'in the field'" (Arthur & Nazroo, 2003, p. 137.), where the field in this context is extended not only to the research process but also to the external events informing the topic.

3.3 Data Analysis

The analysis of data was not focused in one moment in time. On the contrary, it was a continuous process carried out throughout all the research. In this sense it followed the cyclical and iterative process of Action Research, where new data and insights continuously inform the current and future actions in the research, as well as the whole direction of the inquiry process. Moreover, data gathered through different methods was constantly compared. By doing so, further links in the research emerged, which frequently shifted the research goals and interests, re-shaping the initial plan as the research progressed.
The "raw" data gathered through the qualitative data collection methods was in written form. This includes verbatim transcripts resulting from the interviews and field notes resulting from the observations. When the interview was conducted in other languages other than English (Italian or Finnish), the interview was also translated to English.
The first step to untangle such content is, as described by Ritchie, Spencer & O'Connor (2003, p. 221), familiarization. It includes reviewing the material gathered with the goal of finding recurrent themes.
After crafting an initial framework, the data gathered was indexed according to relevant categories. The term "indexing" is used in place of the term "coding", a choice aligned with the work of Ritchie, Spencer & O'Connor (2003) since it underlines the idea of how the data "fits" the categories. The starting point was represented by indexing the interviews, and secondly the data was integrated with material coming from

	OBSERVATIONS		SEMI-STRUCTURED INTERVIEWS				
	Notes	Pictures	Notes	Recording	Transcription	Translation	Pictures
Experts' interviews							
Franscesca Foglieni			√	√	√	√	
Tua Björklund			√	√	√		
Kirsikka Vaajakallio			√	√	√		
Taina Mäkijärvi			√	√	√		
Migrant Youth Helsinki staff							
Project Manager			√	√	√		
Product Owner			√	√	√		
Service Desigenr			√	√	√		
School visit							
New Teacher	√						
Experienced Teacher 1			√	√		√	
Experienced Teacher 2			√	√		√	
10 2nd Grade Students	√	√					
4 6th Grade Students	√	√					

Figure 9. Overview of the data collection methods and the participants.

the field notes. Sometimes, when the questions asked in different interviews were the same, the questions themselves were used as first framework for categories, to compare the interviews. This has been the case especially when analyzing the experts' interviews. Both digital and physical way of indexing the data were used during the analysis. The same analysis method was used also during the literature review, when categorizing the benefits of co-design in service design projects.

3.4 Ethical Considerations

As the research is carried out in close collaboration with actors from the educational context and minors, considerations are needed in order to design the field work in an ethical manner.
According to Denscombe (2009), the main rule to follow for researchers is to ensure the protection of participant's interest. This occurs when the rights of participants are prioritized, the research follows an ethical code, consent from a relevant committee or board is given before starting the research, and when the researcher is open, honest and transparent (Denscombe, 2009, p. 59).
Concerning seeking approval, this master's research is approved by principals of the schools and Helsinki Kaupunki through the release of a

research permit to conduct fieldwork in schools. Acquiring a research permit was a time-consuming activity, that required strong management skills, but it was a crucial point in order to start conducting the research.

As for protection of the interests of the participant in the research, extra attention was directed to children. A high degree of trust was given to the teachers to choose suitable times, dates and participants.

In order to avoid stress and feelings of intrusion, the research team tried to apport as few changes to the environment and the schedule of the class as possible. During the school visit, the field work activities such as observations and brief interviews were carried out with a sensitive attention towards the class environment and the privacy of the children. In this sense, the choice of letting the Finnish research being the main point of contact with the children helped in creating a comfortable environment for them.

Consent forms were sent beforehand to guardians of the children in order for them to give permission to participate in the research. Exhaustive information about the research and the purpose of the data collection was included in the form, as well as the possibility of rising more questions directly to the researchers. Permission to audio record the conversation was given from the participants.

In addition to gather the consents from the guardians, it was crucial to make sure that the children themselves wanted to participate in the research. It was the priority of the researchers that, where possible, children themselves decided about their participation in the research. Considerations on how to create different kinds of possibilities for joining the research were made, for example by giving the possibility to participate alone or in a group.

This ensured that children would feel in a safe and comfortable environment.

Another consideration in obtaining the consent for research from children concerns the fact that they might feel obligated in taking part of the research since it happens in the school environment. Nevertheless, this feeling was mitigated by the non-mandatory nature of the Buddy-school program.

During fieldwork, no children refused to participate, nor presented signs of unwillingness to their participation. In order to protect the identities of the children and teachers taking part in the study, individuals and organizations (namely, the schools) remain anonymous. In fact, no names nor school identities will be disclosed in this master's thesis. As for the pictures, the ones appearing in this master's thesis which show children in the school setting come from the Helsinki public image bank and are therefore open for the public to browse and download.

In addition to considerations on the participants and the methods of engagement, ethical considerations must also be risen concerning the methods used in the field work. Are the methods chosen (observations and brief interviews) child-friendly and school-friendly? Might children be excluded from this activity because of a physical disability or a learning difficulty? Is the fact that the interview is being audio recorded going to affect the child negatively? How can we ensure that the questions proposed to the children are in fact well targeted and do not create negative outcomes?

Moreover, the adaptation of the methods in action will carefully need to follow these considerations and questions.

Between the planning and the execution there usually are several unexpected factors that come into play. Nevertheless, when changing plans and quickly adapting the methods to the real life setting that the research is experiencing, it is important to bear in mind the ethical considerations, not allowing the pressure of the field work to jeopardize the ethical standpoint.

3.5 Methodology´s limitations

While the employment of the methods described in this chapter, represented a fruitful and appropriate way to gather data, it is also crucial to recognize the limitations associated to the methods themselves and to the implementations of such methods in this particular research project.

Firstly, time limitations were strict. Due to the time frame of this master's thesis, a restrict number of interviews and observations could be implemented. This resulted in the not having a vast variety of data, meaning that the findings of this research will be tied to the amount and typologies of participants in the interviews. For example, it would be interesting to conduct investigations which include other actors in the schools, such as principals or administration workers. Moreover, teachers have very busy schedules and it was not easy to contact them. Even when they found time to dedicate to this research, their availability was short in time. This resulted in the need to hold brief interviews, which made the task of going deep in the questions difficult.

Apart from the tight timeframe, language was also a "management" limitation in the employment of research methods.

Although the presence of a Finnish-speaking researcher was beneficial, the preparation for the fieldwork, as well as the analysis, required more time because of the translation from Finnish. In addition to that, language barriers were found also outside of the field visit. This required work from both sides, trying to find a common ground for understanding, sometimes preventing the possibility of having a rich conversation on the topic because of the basic language used. In addition, although the translation from Finnish to English (as well as the translation from Italian to English) was done in the most careful way, it was unavoidable to lose nuances in the translation. This might have impacted the way data was interpreted in English. Moreover, the observations made by the English-speaking researcher in the field visit to the school were carried out without understanding the spoken dialogues between the children and teachers. If on one side this represented an advantage, allowing to focus on the non-spoken interactions, on the other side it could represent a way to misinterpret the situation. This leads to underline how all the results and insights are a result of a personal interpretation and reading of the data, shaped by collaborative work and confrontation with other actors in the research.

Another limitation in data gathering is represented by quick and on-spot adaptations of the original field-work plan. Even though these changes were always apported with a topic guide in mind, they might have affected the data collection and what kind of data was gathered.

Finally, the role of the observer must be addressed. Since two external figures (the researchers) joined the Buddyschool class, the usual environment was necessarily altered. Attention on how the presence of strangers in the class changed the behavior of the children was taken into consideration. Even though the fact that the children were observed did not seem to apport consistent changes in the environment, it cannot be affirmed for sure.

Chapter 04

Literature Review

4.1 What is Co-Design

Before delving into the co-design evaluation issue, the following section provides an overview on what is co-design.

According to Sanders & Stappers (2008), companies have been more open to approaches that involve taking into consideration the opinion of the users. This willingness started with user-centered design, where the user is seen as subject, a passive entity that provides assessments on ideas and prototypes created by others. In contrast to the passive role of customers, during the 1970s Nordic Countries began to implement a more active way of engaging with the users, whom are seen as active contributors, becoming "partners" for the organizations (E. B.-N. Sanders & Stappers, 2008).

These Scandinavian practices evolved to become a more complex living portrait of various disciplines and approaches. This "landscape" fostered the development of concepts such as co-design and co-creation (E. B.-N. Sanders & Stappers, 2008).

In line with the definition given by Sanders & Stappers (2008) this thesis ustilizes the term co-design to express the "collective creativity as it is applied across the whole span of a design process" (E. B.-N. Sanders & Stappers, 2008, p. 6). In contrast, co-creation refers to all type of collective creativity that two or more people share together. So, if co-creation represents thejoint creation and the collaboration between two or more stakeholders, co-design is the actual act through which co-creation happens (E. B.-N. Sanders & Stappers, 2008).

According to Kleinsmann (2006), co-design can be described through three key characteristics. Firstly, it creates knowledge and allows it to be shared across various actors. Secondly, co-design allows participants to communicate clearly about what is the object of the design process and the process itself. Leading to, thirdly, also allowing shared understanding on the same matters (Kleinsmann, 2006, p. 38). In this sense, co-design is defined by the author as "the process in which actors from different disciplines share their knowledge about both the design process and the design content. They do that in order to create understanding on both aspects, to be able to integrate and explore their knowledge and to achieve the larger common objective: the new product to be designed" (Kleinsmann, 2006, p. 38).

If Kleinsmann's view emphasizes generating and sharing knowledge amongst the participants in the co-design process, Steen (2013) argues that a definition of co-design that encompasses the combination of "thinking and feelings, facts and values" (Steen, 2013, p. 28) is needed. He delineates co-design as "a process of collaborative design thinking: a process of joint inquiry and imagination in which diverse people jointly explore and define a problem and jointly develop and evaluate solutions. It is a process in which participants are able to express and share their experiences, to discuss and negotiate their role and interests, and to jointly bring about positive change" (Steen, 2013, pp. 28–29). The author also argues for the importance of unveiling the ethics of co-design, which often are not explicitly stated. This is important in order to foster reflection in the participants. This process can lead to a higher awareness in the participants, who become more conscious about the co-deisgn process and their participation and involvement in it (Steen, 2013). This thesis argues that addressing these ethical questions presented by Steen is a key step in ensuring that co-design participants are aware of the possibilities and the benefits brought about by the process and can be more self-aware on the methods and the extents in which they can engage with the process.

In this sense, co-development in service design projects could be defined as the process of jointly developing a service, sharing and co-creating knowledge, as well as researching the problem and imagining a solution together.

Indeed, accepting to incorporate co-design can be hard for organizations, since it stands as a threat to the established processes that are built

on the hierarchical power structures that characterize some organizations (E. B.-N. Sanders & Stappers, 2008). Nevertheless, change is possible, and it has been happening. According to Sanders & Stappers (2008) this is also thanks to the large utilization of web platforms from users to express their opinions and ideas.
In general this shift from an organization-centric view to a customer-centric approach, but even from a user-centered view (users as passive subjects) to a co-design approach (users as active participants) is underpinned by a fundamental change in the roles played by the actors taking part in the process (E. B.-N. Sanders & Stappers, 2008). According to the authors, co-design seeks to engage the users as "experts of their experiences" through the utilization of tools aimed at sparking ideas and opinions.
If these tools are suitable enough to let the users express their passion and expertise on the topic, users might become co-designers, which means they would actively contribute throughout the whole process. In this setting, the role of the user is not the only one which undergoes major changes. Designers and researchers have to rethink their contribution and approach, too. It is sometimes the case that the designer role and the researcher role are played by the same person, who needs to become familiar with the art of facilitation. Being a mere translator of users' needs is not enough anymore. In co-design, the ability to facilitate the ideation process is needed alongside the skill of being a visual and system thinker, trained to conduct and take part in a design process (E. B.-N. Sanders & Stappers, 2008).
But even if many organizations, both in the private and public sectors, have already decided to shift their focus from designing for their users, to designing with their users, sometimes the term co-design is used just as a circumstantial term, separating it from its meaning and the actions that it requires (Steen et al., 2011). According to Steen et al. (2011), in addition to co-design being used as a buzzword, the contribution of co-design to a project is not always easy and clear to argument for and outline. In fact, enthusiastic studies about the potential of co-design are accompanied by more skeptical point of views that argue an absence of a clear way in determining if the engagement of users actually affects the design process and the results, and how (Ostrom et al., 2015; Trischler et al., 2018). Furthermore, co-design can also just become a trend useful for organizations to push their products or services into an already overcrowded market (E. B.-N. Sanders & Stappers, 2008). Even though the financial drivers are important for an organization that chooses to enter the consumer market for profit reasons, marketing and brand development should not be the only reasons why an organization chooses to employ co-design. Involving users and stakeholders needs to be implemented and managed in the right way, otherwise the risk will be minimizing co-design to become a tool for financial success, without understanding and benefitting properly from its core characteristics. Therefore, understanding the value and the benefits underpinning the co-design process is crucial for every participant, from the users to the employees and managers in the organization, including, of course, designers.
In summary, in a world where co-design is often used as a buzz-word, there is a need for developing a set of methods or tools to measure, showcase and evaluate if the benefits promised by a co-design approach are actually realized (Steen et al., 2011), and if they are actually deriving from the use of co-design in the project. This would help in understanding the various forms of the co-design value and allow to showcase them.
In order to be able to discuss this matter further, an overview of the benefits of co-design is presented in the next section.

4.2 Co-design Benefits

As consumer market is changing, companies are realizing they must abandon their established ways of working and turn from a company-cantered approach to a customer-cantered one.

Abandoning the "firm-centric" point of view includes changing organizations' view on customers, who are increasingly seen and engaged with as collaborators that can co-create unique value together with the company (Prahalad & Ramaswamy, 2004).
This shift in the perception of users suggests that the market is transitioning from a concept of value created and delivered by the company, to a concept where value is co-created between all the relevant stakeholders, including the company and the customers, all participating in a particular way, depending on their access to resources (Vargo & Lusch, 2004). In other words, "the focus is not on the offering, per se, but rather on the customer's value cocreation process" (Ostrom et al., 2010, p. 26). The rising conception of market as a venue of shared creation of values implies a change in the conception of value itself, and the process of value creation. Although recent research explored this shift towards a co-development process (e.g. Sanders, 2002; Sanders & Stappers, 2008; Steen, 2013), there are several issues that need to be investigated more. According to Ostrom et al. (2010), a research priority that needs to be addressed is research on measuring the impact of involving customer communities in the process of service development. For instance, further inquiries are needed regarding ways of measuring co-design methods in order to manage them more properly and accurately. Furthermore, the authors suggest the need for a deeper research in frameworks aimed at measuring the economic and noneconomic benefits of co-design.
One of the reasons why companies are increasingly interested in co-development is the vast array of benefits brought about by such approach. These benefits are usually narrated through success stories of companies that show the impacts that co-development can bring in different areas. Not only the private sector is showing interest in these kinds of methods, but also the public sector is increasingly keen on implementing co-design approaches in order to co-create services that are the most valuable for the citizens (Pirinen, 2016).
While co-design benefits are recognized by practitioners and literature, there seems to be considerably less research on ways of measuring them. This research gap is addressed by both practitioners and researchers, who show a growing willingness for investigating evaluation opportunities aimed at assessing the effects that collaborative approaches are bringing to the projects and services (Steen et al., 2011).
In order to understand the multiple and different reasons for evaluating co-design, it is first essential to shed light on its benefits. In other words, by presenting a brief overview of the results of co-design in service development projects, this section aims at inquiring the typologies of benefits displayed by literature in order to later validate them through the case study. In addition to researching what kinds of benefits are brought about by co-design approaches, the issue of which evidence is needed in order to prove the link between the process and the service results is raised.

Creating and managing useful collaborations is a key skill that companies need to learn in order to have a "collaborative advantage" (Kanter, 1994, p. 1). This advantage proves to be excellent in times distinguished by an increasing rise in competition and constant and considerable changes in the service sector. Interestingly, the effects of co-design are multifaceted and can show benefits on various fronts. Evidence suggests how companies that employ co-development with their customers are more successful than the ones who do not invest their resources in reaching out to customers and inviting them to be part of a co-design group. The effects of customer involvement in development of services include benefits on the outcomes, on the perception of the service and on the customers' attitude towards it (Steen et al., 2011). It also has consequences on the performance of the organization, including employees' satisfaction and financial performances (Ostrom et al., 2010, p. 24). The following section will review relevant literature showcasing the benefits that

co-development might bring to the final product or service, to the company or to the customers. From several researches on the topic, six of them (Alam, 2002; Hoyer et al., 2010; Kristensson, Magnusson & Matthing, 2002; Kujala, 2003; Steen et al., 2011; Suominen & Pöyry-Lassila, 2013) are selected and analysed. The benefits described in the articles are then grouped and clustered in a scheme in order to highlight recurrent themes.

Steen et al. (2011) analysed three cases of co-design in service design projects and clustered the benefits observed into three categories.
The first group is the effects on the design project, for instance increasing knowledge on users' needs or generating better ideas. Secondly, there are benefits for the customers, for example a better fit for users' needs. The last category represents favourable consequences for the organization, for instance fostering collaboration and innovation practices. From the cases analysed by Steen et al. (2011), several benefits emerged. Firstly, the involvement of the users allows the research team to validate their assumptions. In this way the result of the development process will not be founded on unvalidated hypothesis. Allowing users to express their expertise on the matter as "experts of their experiences" (Visser et al., 2005, p. 10) often makes the result more valuable for the users, who recognize their true needs and inputs taken into consideration. This will generate a stronger ownership of the idea and a better acceptance and higher use of the service, consequently improving customer satisfaction.
Moreover, the involvement of users allows the design team's choices to be more easily accountable for. The findings are easier to accept and results are more convincing when presented to other people, as they are based directly on input, opinions and feedback from the users and customers. In addition to having effects on the customer satisfaction and internal acceptance of the ideas, including users in the development process generated more innovative ideas. Users can offer a different point of view on the everyday work of employees, sparking creativity in their everyday practices. For example, by involving children in designing ICT system, creative ideas are generated that might result in new services implementation. From the employees' side, a better service delivery was enhanced through the creation of empathy towards the customers. For instance, in a logistic service case analysed by the authors, employees were able to understand the pain points and issues faced by customers after having insights concerning the customers' experience. This resulted in a higher motivation to change and deliver faster and improved services. The willingness to change was also translated on an organizational level by fostering cross-functional collaborations in the organizations (Steen et al., 2011).

Suominen & Pöyry-Lassila (2013) analysed three cases where organizations participated in co-design interventions. The case follow-ups were conducted through interviews and the aim was studying the effects of the service co-development. Two out of three cases were set in the school system. Interestingly, the benefits emerged during the follow-up interviews carried out by the authors are linked to both a practical and a mindset level. The authors underlined how involving users in the development process sparked changes in the thinking and mindset of the users and of the organizations, as well as fostered change in the actions undertook by the company. For instance, a change observed in the way participants were thinking is highlighted by the fact that after joining the co-design process, they were able to understand the co-creative nature of value and see the importance of empowering different stakeholders and users. Although this research suggests that awareness on the process is created through experience, it would be interesting to compare the gained level of process awareness between the participants coming from the company side and the users. Nevertheless, the change in mindset from the company side can be manifested in the

increasing cross-functional and process thinking, which fosters new collaborations from different departments of the organization.
In addition to that, the interviews show how new internal and external networks and collaborations were initiated, as well as knowledge transfer within the organization. Participants were keen on sharing valuable knowledge with other people in the organization who did not participate in the co-design process, allowing further interaction and collaboration to take place. Evidence of successful knowledge transfer was also found in the willingness to further utilize the materials produced during the process. In one case analysed by the authors, the knowledge and materials produced were used in further development sessions inside the organization, who tweaked the material accordingly to their new needs (Suominen & Pöyry-Lassila, 2013).

Kujala (2003), reviews three streams of research based on ICT system design in order to showcase the benefits of user involvement. Firstly, when involving the users in the development process, users' requirements result more accurate. Furthermore, the efficiency will be higher, as the final result will not comprehend features that the users do not need, since their opinion and feedback was taken into consideration early on in the process. This also results in the users to accept and understand the systems in a more effective way, because they are developed and tailored on their experience and knowledge. Lastly, Kujala reports how users' satisfaction increases, as well as the engagement in the organization's decision-making process (Kujala, 2003).

Additionally, Hoyer et al. (2010) categorize the positive outcomes of implementing co-creation in new product development projects in two clusters. On one hand, co-design improves efficiency because of cost savings in various development steps (from reduced risk of product failure, to not having to pay for ideas generation), which have a direct influence on the performance of the organization, increasing employee satisfaction and eventually profitability. On the other hand, co-design improves effectiveness because the market fit is more appropriate, and the product or service has more commercial attractiveness. The benefits, as found by Hoyer et al. (2010) include similar effects showcased by the previous authors. For example, the authors describe how products or services created with the engagement of users have a higher effectiveness, since they have a closer fit to the customers' needs, and they are perceived as of higher quality and novelty. In addition to that, costs are lower as a result of the pro-bono input of the customers and a faster speed-to-market, including a lower risk. Cost reductions are seen also in the lower marketing costs. This is because less is needed in order to market the product or service, since they will spread to the market faster because of word of mouth as they are more valuable for customers. Finally, financial savings are also represented by a lower customer education and support cost, also due to the fact that potential issues with the product or service will be spotted early on in the process (Hoyer et al., 2010).

Moreover, Kristensson, Magnusson & Matthing (2002) conducted a study aimed at researching if including customers in the development process would lead to more original ideas. During their study, they found that involving users leads to generating more creative ideas because customers are not constrained by technology and specialist knowledge coming from the industry. Users do not hold the burden of technology-related constraints, and they can focus on proposing solutions that integrate new and original ideas more freely in their everyday contexts. Moreover, the authors found that the end result will fit better with actual users' needs, since costumers participating in the development process offer insights on their everyday life and conditions. This translates to a more effective identification of the users' requirements, which in turn leads to success in the product or service (Kristensson et al., 2002).

Lastly, Alam (2012) performed exploratory interviews with practitioners working with cases of service design, in order to identify elements of user involvement. The author found that co-design leads to a higher quality of the service delivery. In fact, the service is more differentiated and serves the users and customers in a more meaningful way, resulting as a more valuable service. In addition, they noticed how service diffusion and market acceptance are faster, resulting in easier and better education for the users, as well as a quicker support service. Relationships with the customers are improved on the longer run, meaning that a continuous collaboration is fostered. Finally, the cycle time of development is lower because issues can rise early on and they can be addressed more efficiently (Alam, 2002).

It is interesting to underline how several of the effects presented by the various authors overlap. Apart from mentions about lower costs and financial benefits, the consequence that all authors seemed to have noticed in their research is the increased customer satisfaction. The following table reports a brief summary of the benefits individuated in the six studies taken into consideration.

From:	**Reported Benefits of involving users in the service development process:**
Steen et al. (2011)	1. Result of the development process is validated by users 2. The result is valuable for users 3. Choices are accountable for because based on customer opinions 4. Innovative ideas 5. Creativity amongst employees can be fostered 6. Customer satisfaction improves 7. Better service delivery 8. Willingness to organizational change
Suominen & Pöyry-Lassila (2013)	1. Participants understand the value of the process 2. Increase in cross-functional and process thinking 3. New internal and external collaborations 4. Materials and knowledge from the co-design setting are used again in the organization 5. Knowledge transfer within the organization
Kujala (2003)	1. Accurate users' requirements 2. High efficiency 3. Better acceptance and understanding from the users' side 4. Improved satisfaction 5. Increased engagement in the organization's decision-making process
Hoyer et al. (2010)	1. Reduction in costs of generating ideas 2. Faster to-market process 3. Higher effectiveness of services or products 4. Success of product or service is more likely 5. Less advertisement needed 6. Lower customer education and support costs 7. Potential issues are raised early in the process 8. Stronger customer-firm relationship
Kristensson, Magnusson & Matthing (2002)	1. More creative ideas 2. Better fit in users' needs 3. Better users' requirements
Alam (2012)	1. Higher service quality 2. Lower cycle time 3. Users' education is better and easier to achieve 4. Service diffusion and market acceptance are faster 5. Quicker support service 6. Improved relationships with customers

Table 1. Summary of benefits of involving users in the development process of services based on the literature review.

In addition to the categorization by author, the co-design benefits presented in the six studies mentioned in this section can also be clustered according to overarching themes. *Figure 10* shows how the effects identified by the different studies can be grouped according to general benefits, represented by the top boxes. The general overarching benefits identified are: reduction in costs, higher service quality, higher users' satisfaction, better ideas, fostering collaboration, and changes in mental models. Some of the specific benefits individuated by the literature review could belong to more than one cluster. For example, the fact that service diffusion is faster leads to decreasing costs but is also symptomatic of a higher satisfaction level from the users.

Moreover, the overarching themes influence one another very closely. For instance, the fact that co-design leads to better ideas will likely result in a higher quality of the service, where the value is strongly perceived by the customers. Furthermore, higher quality and higher satisfaction closely influence each other, affecting in turn the level of financial inputs required by the project.

It is interesting to observe how the benefits presented here can be applied to different actors in the process, as also discussed by Steen et al. (2011), who cluster the benefits into the ones aimed at the service's customers, the ones aimed at the organization, and the ones aimed at the service project. This is applicable to the categories featured here because each general theme can be adapted to different actors. For instance, a mindset change can occur both internal to the company or within the users participating in the process. A shift in perspective from the company could potentially bring an impact on the internal processes, skills and new capabilities, whereas a mindset change in the users might result in a more open-minded approach towards co-design and the service itself.

In general, it is clear how co-design implementation in the service development process brings about a wide array of benefits, which space from financial benefits, to improved customer satisfaction, not forgetting potential benefits concerning a change in the mindset of the participants in the process. In this sense, participants are more open to the concept of participation on several levels (internal, external, with customers) and seem to understand the value of the co-creative nature of co-design process.

This brief literature review underlines how there are expected and potential benefits as a result of the right application and management of co-development processes inside organizations. Nevertheless, these effects are not fully measured, yet. In addition to offering an overview of benefits, literature also suggest the need of further inquiry into metrics or indicators capable of measuring the impacts of co-development.

Steen et al. (2011) highlight the "need for developing ways to monitor and evaluate whether the intended benefits are indeed realized" (Steen et al., 2011, p. 53), and also to which extent. Hoyer et al. (2010) underline how it is now necessary to define metrics and measured of "economic and noneconomic benefits of consumer cocreation" (Hoyer et al., 2010, p. 292), asking themselves "how can a firm measure the benefits of the co-creation process?" (Hoyer et al., 2010, p. 292). Similarly, Ostrom et al. (2010, p. 25) address the need to manage and measure the impact of involving consumers in the development process.

In conclusion, there is a need for further research in the area of evaluating co-design, including assessing if the co-design process brings about benefits that are measurable in the final outcome and can be attributed to the implementation of a collaborative development of the service.

In the next section, an overview of the evaluation discourse within the service design practice is offered in order to gather further comprehensive knowledge on the matter.

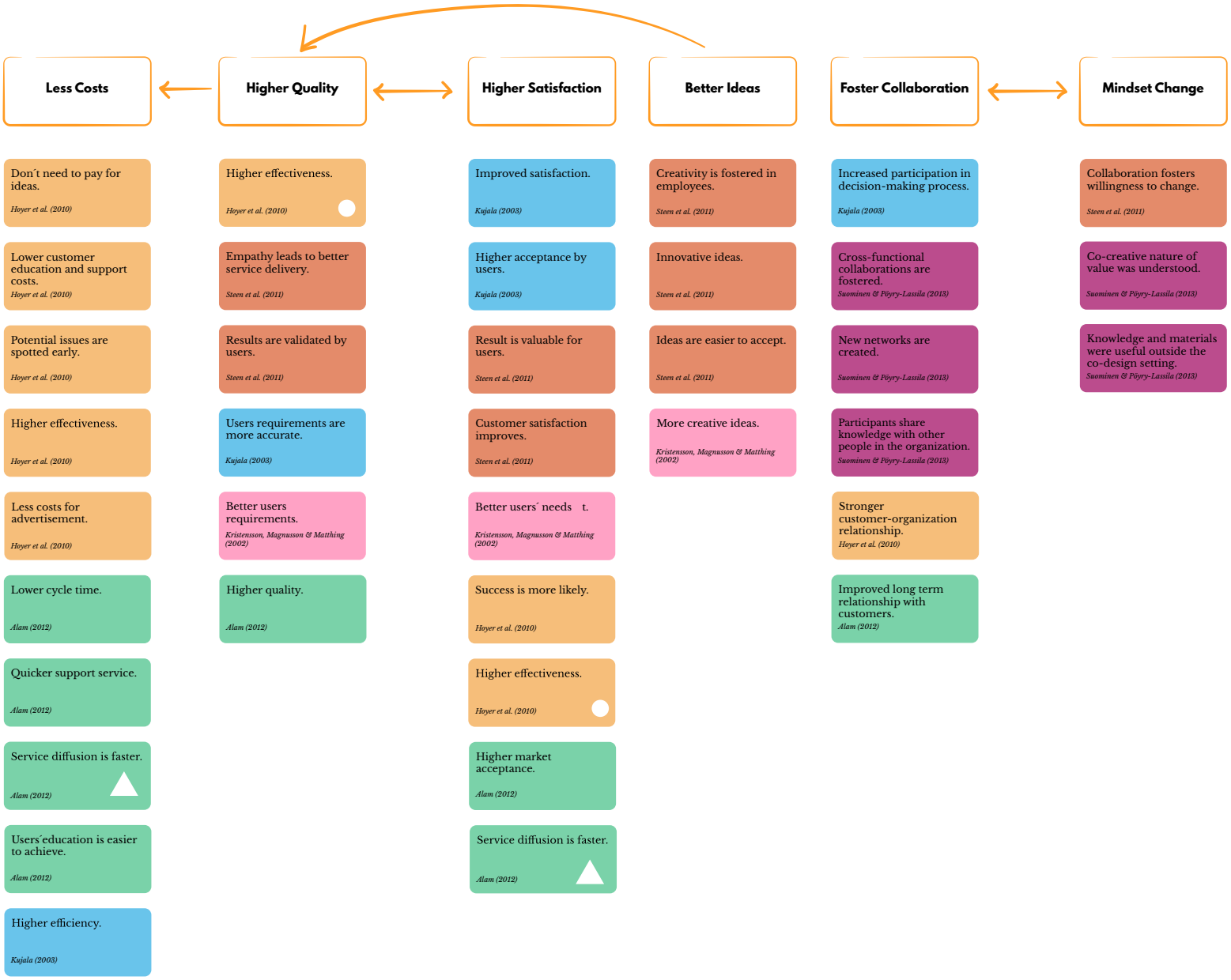


Figure 10. Categories of co-design benefits in service design projects.

4.3 Evaluation of Co-Design in Service Design

In general, evaluation can be defined as the "process of determining merit, worth, or value of things – or to the result of that process" (Scriven, 1991). In this sense evaluation can be the process of evaluating specific characteristics and matters, but it can also refer to the product of the act of evaluating, for example a report (Scriven, 1991).

Evaluation as a field comprehends different topics and groups that apply the theory of evaluation in different areas, such as product evaluation, performance evaluation, program evaluation, or personnel evaluation (Scriven, 1991).

From the 1970s onwards, a specific stream of research was set up by scholars interested in studying evaluation theories, methods and practices (Coryn et al., 2017). This stream is called Research on Evaluation (RoE) and its aim is analyzing existing practices and knowledge on evaluation, in addition to contributing by generating new theories and understanding of evaluation practices.

When it comes to research concerning design and evaluation, it is still fragmented and far from being comprehensive. Nevertheless, the topic has been recently gathering more and more attention, as demonstrated by the several recent blog articles, conference proceedings and peer-reviewed articles that combine evaluation and design.

In general, when talking about evaluation in design there is not a set of common terms or practices, and even the topic itself is still being explored. For example, when articulating evaluation within the service design sector, what is actually being taken into consideration? Is the goal to evaluate service design practice or the service design project? Is the evaluation targeting specific components of the process and their implication on the results as a whole? Foglieni, Villari & Maffei (2018) express similar dilemmas and difficulties on framing evaluation within service design practice. These obstacles were also found in the framing of this research.

Even if a specific approach and glossary that is shared amongst practitioners and academia is not formed yet, there is a common interest in exploring more around these issues and talk about these topics. This might be partially linked to the willingness of improving one's own practice and learn from it, but also perhaps for the need of constantly advocating for design by showing meaningful data that can support the choice of employing a design approach (and get funding, or a job). A more in-depth discussion about the motivations for evaluating design will be featured in the next chapter, *5. Analysis & Findings*.

According to Foglieni, Villari & Maffei (2018), there are two possible paths in evaluation within service design practice. One option is evaluating services and the other option is evaluating service design. They argue how the first, evaluating services, is easier and more approachable than the latter, evaluating service design. In fact, their work focuses on studying service evaluation, for which they also created a framework and a set of guidelines (Foglieni, Villari & Maffei, 2018).

One of the most interesting part of their research concerns the theorization of an "evaluative research", which involves the use of design research tools (such as observations, interviews, surveys, and focus groups) with the goal of gathering data aimed at expressing opinions and judgments in relation to the key values that were identified at the beginning of the project. In fact, according to the authors, evaluation means setting goals in terms of value and see if the service (or concept, prototype, service design itself) achieves those values (Foglieni, Villari & Maffei, 2018). This definition highlights how evaluation is closely tied with values, meaning that evaluation needs to encompass both empirical and normative aspects, since it is indeed the value component that differs evaluation from other kinds of inquiries (Foglieni & Holmlid, 2017).

Part of the reason why it is so difficult to articulate measurement discourses in service design

practice is because of the difficulty to isolate the features to assess. Service design processes and outcomes are always situated in a context and they need to be understood in that peculiar environment. The context consists not only of the physical environment and surroundings, but includes also the actors and the social interactions, both towards other customers and towards the facilitators of the service (Gupta & Vajic, 2000). These actions and the set of social rules that can be found behind each behavior, representing the main lenses through which the actors interpret the experience and the service (Gupta & Vajic, 2000). This is why isolating the outcomes or the process from the context in which they take place is very difficult.

On the other hand, although implementing a strict set of measures and well-defined indicators might be hard because of the dependence to the context, crafting rigorous frameworks for evaluating, comprehensive of clear and precise metrics, is also needed. This would make it easier to have comparisons and would be more relatable for professionals who speak a "different language", such as managers or engineers.

This dichotomy between contextual inquiries and precise measurements can also be found in the interviews with the experts, presented in the next chapter, *5. Analysis & Findings*.

Concerning the evaluation of participation, Burton (2009) highlights how the two most common methodologies are based on constructivism and case-study methods. On the other hand, according to the author, more quantitative-oriented approached and experimental designs are becoming more and more common. Carpini et al. (2004) argue that the best approach to measure citizen engagement would be combining the strong points that comes from multiple methods and methodologies, not forgetting that the context is actually the factor that affects participation the most.

Instead of arguing for either qualitative or quantitative approaches to evaluation, a way would be considering both approaches, since both offer useful point of views, methods and techniques that can benefit the research. The synthesis can be represented by the realist approach described by Pawson and Tilley (1997). This approach combines the importance of the context with the possibility to identifying "underlying" casual mechanisms. In this way, for example, anecdotal and practice stories are as valued as more empirically and experimental driven methods (Burton, 2009).

4.4 Evaluating the Impact of Co-Design

In addition to understanding the discourses around evaluation within the service design practice, it is also valuable to look at the motivations behind impact evaluation. Understanding the motivations that lead designers to talk about impact evaluations is useful in order to understand better the kinds of indicators that should be set into place. In other words, understanding the why will be helpful in order to frame the how.

In 2017, the city of New York hosted the Measured Summit, a conference where experts in different fields of design came together in order to discuss the topic of evaluation and measurement. During the conference, a distinction between measuring the impact in an evaluative way and in a formative way was made. Such distinction, already part of the work of Chess (2000) and cited by Burton (2009), is presented in an article by Cat Drew (2019). She highlights how impact has traditionally been evaluated following an evaluative approach, meaning by focusing on proving that something has worked or that it will work, but as designers we must also focus on measure the impact in a formative way, thus focusing on the learning experience of the process to understand how to improve (Drew, 2019). According to Drew (2019), organization might be so much focused on measuring what is crucial for them that they forget to assess what users' value. This is where the intuition of experienced practitioners might be helpful in order to direct the evaluation activities (Drew, 2019).

Therefore, both private and public sectors are seeking the expertise of this kind of professionals, wishing for them to help assessing and showcasing the impact that the organization's design activities had.
City of Helsinki is no foreigner to this discussion, since recently it partnered with the Department of Design at Aalto University to address similar challenges (Helsingin Kaupunki, 2019). The need to evaluate the impact that the design activities had within the city sparked from the desire of new tools for accountability and communication. From six different design categories which provides a summary of the wide landscape of design activities in the city of Helsinki, tools and metrics were created to measure the impact of design in specific projects (Helsingin Kaupunki, 2019). The fact that every design project taken into consideration in this specific project by the City of Helsinki had different indexes in order to showcase the impact of design acts as a reminder on the need to consider how different design fields, different design processes, and even different design projects might have to be evaluated in different ways, with their own set of metrics, methods and characteristics (Heller, 2017). What argued by Heller (2017) proves to be true if the previous considerations on how evaluation is context dependent are taken into account. On the other hand, though, this does not exclude the possibility for the learnings to be applied to other cases.
A way in which impact measurements should be different in various organizations or projects is, for example, according to their maturity level (Björklund, Hannukainen, & Manninen, 2018). According to the authors, different maturity levels, visualized by the Design Ladder (Danish Design Centre, 2001), require different tools and strategies to evaluate the impact of design because the Key Performance Indicators (KPI) to measure will be different. For instance, if one organization has reached the third level of the ladder, "design as process", it means that the design practice is well established within the organization and has become an integrated part of the ways of working. Accordingly, measurements that take into consideration what customers think and value become more important, in contrast with implementing just mere performance indicators (Björklund, Hannukainen, & Manninen, 2018). On the other hand, if an organization finds itself on the first level, "non-design", it means that measuring the impact of design is impossible, since there are no or very few design activities in action. On this level, though, it is possible to look at external benchmarks that offer examples of positive impacts (Björklund, Hannukainen, & Manninen, 2018).
Besides reflecting on how measuring the impact of co-design might require different methods and approaches in different organization, the question of whether it is necessary in the first place must be addressed, too. Is a framework consisting of specific metrics and indicators aimed at measuring involvement and participation really needed? The literature and the experts' interviews presented in this research seem to support an affirmative answer. Even though personal experience in a co-design process acts as the best "convincer", metrics and indicators could be useful to draw project managers (for example) closer to the idea of implementing a co-development process. In addition, the literature shows that even if co-design benefits are subjects to several researches and there is an underlying assumption that involving users will generate better results, the research on the impacts of co-design is still fragmented and in need of further inquiries (Steen et al., 2011).
In order to move the first steps into understanding if it is possible to create a framework to measure the impacts of co-design, the next section will explore some of the current metrics used in the service design field, scouting for indicators related to the collaborative performance.

4.5 Metrics used in Service Design

Metrics, indicators, and parameters are certainly not missing from design practice and research.

In a world where measuring and keeping track of data is the key to be accountable and show results and outcomes, the design field is not falling behind. Several metrics used often by organizations are borrowed and adapted from other areas, such as business and marketing. But what are the common metrics and indicators used in design?

Some of the most common metrics used by companies are the ones related to Customer Experience (CX). The utilization of these kind of indicators are spread throughout different organizations in different fields and aim at understanding the rating of the experience of the company, service, or product as perceived by the customer by asking to place a numerical value to it.

For example, the NPS (Net Promoter Score) measures the loyalty of customers by tracking how likely it is that they would recommend the company or their product or services to other people. The NPS is calculated by subtracting who would definitely recommend the company/product/service (scores 9 to 10) from who would not do so (scores 0 to 6). Other examples of such metrics are CSAT (Customer Satisfaction) through which is possible to measure how satisfied the customers are by asking to rate the goods or service experiences, and CES (Customer Effort Score) which tracks how easy and satisfactory it was for customers to interact with the company (Usertesting, 2018). It is essential to highlight how easy it can be for companies to misuse and misinterpret such kinds of metrics.

A rate from 0 to 10 might be reflecting the number of people who are not satisfied with the customer service, or would not recommend the products to anyone, but they fail to depict the reasons behind those choices. By just stopping at gathering numbers on how many clients are happy and how many are not, it is very hard to understand how to improve or what to maintain at the same level. So, even though these kinds of metrics highlight the importance of hearing from the customers and adopting a customer centric vision, it is also crucial to understand that other kind of feedback is necessary to know why and how to act upon the numbers gathered.

Customer experience can be assessed also through SERVQUAL (Service Quality Model). It comprehends metrics useful to analyze the difference between the experience that users were expecting to have and the real experience of the service, according to five (originally ten) determinant categories: reliability, responsiveness, assurance, tangibles and empathy (Adil, 2013; Parasuraman et al., 1985). SERVPERF offers a modified version where quality is not based on satisfaction, but it is seen as an attitude.

But if these metrics come more from a marketing and financial perspective, there are also methods that allow organizations that are not clearly financial-driven, such as NGOs or NPOs, but also companies that want to evaluate with a different focus target than the conventional one, to assess different kinds of impacts, that might also not be monetized. An example is the SROI (Social Return on Investment), which measures values such as social or environmental values who are sometimes not part of the standard assessment carried out by organizations. In addition to opening up the spectrum of possibilities on what to evaluate, calculating SROI also requires the organization to be in touch with their stakeholders, opening up a dialogue that can foster a common understanding and the co-creation of values (Banke-Thomas et al., 2015).

On the other hand, there have also been attempts to measure design activities also from within the practice. An example of this is the Design Scoreboard (Moultrie & Livesey, 2009), which presents a framework where nations can be ranked according to their national level of design. This value is considered on an absolute level of design capabilities and on a relative level of design capabilities. The Design Scoreboard proposes to keep track of enabling conditions, inputs, outputs, and outcomes of design through various indicators, such as public investment in design, number of design graduates, trademark registrations, or design firms. The report is comprising of an overview of the issues concerning the data collection. In fact, it

must be highlighted how these kinds of indicators are often facing a lack of reliable data sources, for example because they are not complete, or they date back too many years. In addition, it is difficult to compare different countries that are very much different in size, for example comparing the whole USA with a smaller country such as Singapore.

Other assessment tools that can be found under a so-called "designerly" way of doing are, for example, Impact Mapping and Storyboards. Impact Mapping is a collaborative mind-map that visualizes the impacts and other key aspects to take into consideration when assessing the impacts and the outcomes. The center represents the goal to achieve and the subsequent layers unfold the people that can help in achieving that goal, the impacts and the way the actors´ behaviour should change and finally the deliverables, exploring how they can support the required impacts (Adzic et al., 2014)

This short list of examples on metrics and evaluation tools underlined some limitations, such as the difficulty to gather reliable data and compare it with data from other sources, as well as the necessity to carry out additional feedback inquiry to interpret the data and uncover the underlying reasons and motivations. In addition, it is important to take into consideration different kinds of impacts and approaches to them.

On the other hand, though, it is noticeable how these metrics foster user centeredness. Adding to that, an interesting argument, that was briefly touched upon in the previous section, is made by Björklund et al. (2018), when they argue that metrics are relevant to different organizations depending on the maturity level in which the organizations are. They cite the Design Ladder, where organizations that do not use design are placed at the bottom. On the second step there are organizations that treat design as a finetuning activity, on the third level organizations that have design embedded in the process can be found, and, finally, on the fourth and upper level design is an integral part of the organization's strategy. This division underlines how metrics and evaluating design activities within the company closely depends on the level of maturity. For instance, in the first level ("no design") evaluating is impossible since there is no design activity present. At this stage it might be worth to utilize benchmarking to look at projects were design is working in an effective way. On the second level, the metrics used might be more focused on the financial return of design

		Cons	Pros
CX (Customer Experience) metrics. *Rate the company's/service's/product's experience*	**NPS (Net Promoter Score)** *Would you recommend it to your friends?* **CSAT (Customer Satisfaction)** *How satisfied are you?* **CES (Customer Effort Score)** *How easy was it to interact with the company?* **SERVQUAL/SERVPERF** *Is the experience you receive the same you expect?*	*They are easy to misinterpret if further research on the reasons behind the results are not carried out.*	*They promote a user-centric vision because they foster the idea that the customers are importnat actors, whom should be addressed in the feedback process.*
Metrics that can also not be monetized.	**SROI (Social Return On Investement)** *Values that might not be part of the standard assessment carried out by organizations.*	*The framework for measuring the social value might be complicated to implement for some organizations.*	*Provides a tool to organizations that are not focused on financial measurements to assess other kinds of impacts.*
"Designerly" way of doing metrics.	**Design Scoreboard** *A ranking of design's national level for different nations.* **Impact Mapping / Storyboards** *Visualizations of aspects to take into consideration when evaluating impacts.*	*It is difficult to compare countries with very different size. The indicators are sometimes not reliable or based on old data.* *It can be a very different "language" for other actors in the process.*	*It offers an overview of the state of design in a specific country.* *It provides a visualization of the process and the important aspects.*

Table 2. An overview of the metrics presented in section 4.5 Metrics used in service design.

activites and in showcasing the experience of customers, whereas on the third level the focus might be more towards measuring the effectiveness of internal process that utilize design approaches. Finally, on the fourth level it might be harder to isolate the effects of design since it is so much embedded in the strategy. Therefore, it is harder to evaluate when the organization is belonging to the most mature level. This conversation adds interestingly to the topic of measuring co-design, since it will be possible to measure co-design only in the settings where co-design has taken place, collocating the companies or the projects at a specific maturity level that might coincide with the highest steps on the ladder. These highest steps, nevertheless, are also the ones where understanding how to evaluate co-design will be more complicated.

Overall, it is interesting to notice how there are multiple options available for tracking very different and a diverse range of indicators and parameters within the service design practice, but there seems to be a lack of measurements concerning collaborative performances and co-design impacts.

Only a few examples on how to measure collaboration in the design process are present in the literature of Design Performance Measurement (DPM). Nevertheless, the majority of the literature concerning performance measurement is still focused on highlighting efficiency as one of the key parameter to take into consideration in order to evaluate design performances. These kind of metrics with a strong accent on the business and financial side often fail to accurately portrait the growing reality and interest of designers to work with hybrid, public and not-for-profit sectors. Even if efficiency is still considered a key indicator in the public sector due to accountability reasons, emphasizing such traits by leaving aside the core values of the program or the organization would be a considerable oversight (Minassians, 2015).

One of DPM study carried out by Yin, Qin & Holland (2009) highlights, amongst other, collaboration as one of the criteria to take into consideration when measuring the co-design process in order to improve it. In addition to mentioning collaboration, the authors argue how two types of metrics can be found: the ones which aim is to measure the outcome of the design process and the ones which goal is to measure the process. At the moment, a lot of indicators belong to the first category since they can be measured only after the product is launched to the market, such as customer satisfaction or ROIs. The authors highlight the importance of tracking other kinds of indicators that could be useful to produce an overview of the process to be able to

	Most important. →				Less important.
EFFICIENCY.	*Decision-making efficiency.*	*Problem solving.*	*Personal motivation.*	*Ability to work under pressure.*	*R&D process well planned.*
EFFECTIVENESS.	*Delivering to the brief.P*	*ersonally responsible/ work ownership.*	*Understand design rationale.*	*Fast and detailed feedback.*	*Managing mistakes.*
COLLABORATION.	**Clear team goal/objectives.**	**Information sharing.**	**Communication quality.**	**Cross-functional collaborations.**	**Shared problem-solving.**
MANAGEMENT SKILLS.	*Decision making.*	*De ne/fully understand roles and responsabilities.*	*Build high morale within the team.*	*Con ict management.*	*Monitor/evaluate team performance.*
INNOVATION.	*Competitive advantage.*	*Select the right creativity concept to implementation.*	*Products lead to future opportunities.*	*High quality product design.*	*Perceived value.*

Table 3. Visualization of the metrics described in the work by Yin et al. (2011). The metrics aimed at measuring collaboration are highlighted.

influence the current project, as opposed to just gather data after the launch and apply learnings to further projects (Yin, Qin & Holland, 2009). A very similar view is expressed also in the work of Voss (1992).
Yin, Qin & Holland (2009) argue for five criteria composing a framework for measuring design performance in order to improve collaborative design. These are efficiency, effectiveness, collaboration, design management, and innovation. Through a questionnaire answered by practitioners, the author could compile a list of criteria useful for measuring the collaboration performance. The most common one was having clear objectives and aims amongst the whole team. Other criteria are the way of communicating and the environment created by the team, the satisfaction of the team and their ability to make compromises, cross-functional collaborations, sharing information and disseminating knowledge, as well as cooperating with other team members (Yin, Qin & Holland, 2009). The research could still be developed further. For example, after establishing the criteria, it would be interesting to see which kind of indicators or measurements can be used to measure the criteria. For instance, how to measure the knowledge transfer? Or the dissemination of knowledge? Or the communication environment? In addition to that, a development point could be taking into consideration different perspectives on the criteria. The literature review showed how there are several receivers of the benefits of co-design (Steen et al., 2013), and it would be valuable to explore the metrics from different points of view, not only from the team perspective. For example, how is dissemination of knowledge different in a team composed by just employees in contrast to a composition which includes users?
In conclusion, measuring the impact of involving users in the development process, both on the outcomes and on the process itself, is still an area in need of further research. Part of this inquiry might focus on looking for and implementing metrics and criteria in order to measure collaborative performance.

4.6 Literature Review Conclusions

Even if co-design practices are on the rise in the design landscape and are implemented more and more frequently in both the public and private sector, research is lacking on ways to measure co-design. In fact, the discourses around evaluation in service and co-design focus mostly on evaluating the service in itself and not the process. Even though service evaluation is an essential practice that allows the assessments of concepts, ideas, and implemented services, there is an evident need for research on service co-design and service design evaluation. Additionally, it is essential to understand the benefits realized by co-design in order to research their impacts on the service results and the process. Previous literature has focused on showcasing the benefits that a co-design process might bring to a service design project, as explained in section *4.2 Co-design benefits*. *Figure 10* groups such benefits according to high level thematic areas. The literature review presented in this chapter revealed several common benefits that are observed in multiple studies. Overall, six overarching themes can be individuated: less costs, higher quality of the service, higher users´ satisfaction, better ideas, fostering collaboration, and mindset change.
Another common point emerged from the studies reviewed is the need for further research on evaluating the impacts of such benefits and the effects that deploying a co-design approach might bring to the final results. In addition to further investigating service design evaluation, in the terminology of Foglieni, Villari & Maffei (2018), elements that require more attention are process-related metrics. In fact, evaluation applied within the service design field cannot be seen only with an evaluative approach, but must encompass also a formative goal (Drew, 2019). This allows designers to focus on the learning experience throughout all the process, underst-

anding not only what to improve, but also reflecting on how to accomplish such task. The idea of "measuring for improving" also stands at the core of promoting process-centered metrics. In this way, the learnings acquired after the end of the project will be still useful for future applications, but changes could also be made *in itinere* thanks to process-centric metrics that would allow the team to assess the collaborative performance and the co-design process as it develops.

In conclusion, the literature review suggests a need for further investigating the topic of service design evaluation, as well as exploring different metrics that could be used during the process and after the service implementation. These metrics would be focused on understanding and assessing the collaborative performance and the effects that the co-design process had on the service results and the people involved. In addition, there seems to be a need of a systematic organization of the knowledge produced so far on this topic in order to collect and compile a comprehensive list of possible metrics and methods.

In the next chapter, the findings from the fieldwork in the case study and the experts' interviews' results will be presented in order to search for a validation, or lack thereof, of the findings highlighted through the literature review.

Chapter 05

Analysis & Findings

This chapter presentes the analysis of the three main fieldworks featured in this thesis: experts´ interviews, school visit, and interviews with Migrant Youth Helsinki staff. For further clarification on the use of the term "service design" and "co-design", refer back to the first chapter, section *1.5 Research Scope*.

5.1 Experts´ Interviews Analysis

In order to gain an overview of the current discourses around evaluation in service design and co-design, four interviews were carried out with service design experts. The interviews were realized in parallel with the literature review, allowing to compare and contrast the respective findings. This comparison is reflected also in the chapters of this thesis, as the topics investigated in the following sections mirror some of the chapters belonging to the literature review, offering the point of views of the four experts. The first three interviews were exploratory ones, aimed at discussing the topic of evaluation in service design and forming a background knowledge on the topic. The fourth interview was a validation interview, conducted with the intention of corroborating the findings from the previous interviews. Three interviews were conducted in English and one in italian.

Every interview was transcribed (the one conducted in Italian was also translated), and later they were analyzed through an "indexing" method, as explained in chapter *3. Methodology and Methods*.
Here, a brief presentation of the experts interviewed is displayed:

Franscesca Foglieni. She is PhD in Design and research fellow at the Design Department of Politecnico di Milano. She is interested in studying how to integrate service design with the evaluation culture, on a theoretical but also pragmatical level. She is general coordinator of the "Master in Service Design" at Politecnico di Milano where, at least for the past three years, she has included a module on the evaluation of services. She co-wrote the book "Designing better services. A strategic approach from design to evaluation".

Tua Björklund. She is a professor of practice at Aalto Design Factory. She has a background in Cognitive Science, focusing on how people perceive and process information. She is a PhD in Organizational Learning. She is one of the co-founder of Design Factory, where she researches development of design in organizations and how to advocate for design within those organizations. She is co-author of the paper "Measuring the Impact of Design, Service Design and Design Thinking in Organizations on Different Maturity Levels".

Kirsikka Vaajakallio. She is a PhD in Design at the Department of Design at Aalto University. She works as lead service designer at Hellon, a service design agency in Helsinki. Her work is focused on raising awareness on human-centered design and helping large and small organizations to transition towards a more customer-centric way of working. She was involved in several co-design projects, such as the ATLAS project, which dealt, amongst other topics, with understanding the circumstances that facilitated design impact and the obstacles and barriers to it.

Taina Mäkijärvi. She is a Lead Business Developer at Nordea Global Private Banking. She is also a PhD Student at the University of Eastern Finland at the School of Economics' Innovation Management Department. In her work, she is interested in developing customer service concepts, as well as understanding how service design can change the organizational culture towards a more customer centric approach.

In the following sections, an overview of the topics discussed in the interviews and the respective findings are presented.
During the interviews, both service design

evaluation and co-design evaluation where addressed. Discussing the general area of service design evaluation helped in getting the conversation started for later on (during the interviews or during analysis) scoping down to the aim of this master´s thesis, co-design evaluation. Moreover, during the interviews, co-design was treated as an integral part of the service development process: the co-development, iterative and empathic components of co-design were stressed.

5.1.i Evaluation in the Service Design Practice

1. Lack of common terminology

As suggested by the literature, the topic of evaluation and impact evaluation in service design practice is still understudied. While the topic of evaluation in the service design practice is increasingly gaining attention from the research and practice worlds, it is still characterized by a lack of common terminology and shared definitions. During the interviews, the terms used were evaluating, showcasing, measuring, and assessing. The interviews highlighted how discussing about the topic might be challenging due to the lack of common terminology. In fact, "evaluating" can bear different nuances. Nevertheless, one point that was common amongst the experts was associating measuring with an action that requires to be exact and precise, often having numerical results.

Evaluation was still considered a formal process that focuses on indicators, but on a wider level. Showcasing was treated as a less formal process, aimed at investigating the different typologies of impacts.

"To me evaluation would mean that there is a formal process, typically something that a company wants and that is more institutionalized, whereas our work [interview studies with companies] is more focused on capturing where you can see the impact and in what forms and how could we communicate the impact. That of course could be complementary, but it's a bit different. So, to me evaluation is more like a formal assessment and I wouldn't say that I have been part of "evaluating" in these companies… assessing yes, but not really an evaluation process." – Tua Björklund

"I think it's important that we are able to show the diversity of the impact. We can create a framework or picture where we show that when we do co-design, for example, throughout the all organization it has different levels of impacts." – Kirsikka Vaajakallio

These multiple nuances in evaluation are crucial to describe a wide spectrum of motivations and means of carrying out an evaluation, but they could also represent an obstacle, if the understanding around them is not shared and formalized. Another difference that needed to be clarified during the interviews was the difference between impacts, outcomes and results. The interviewees seemed to agree on:

"The impact would be broader, it would include the side effects, the more process-related effects. Whereas the result output would just be kind explicit project goal or outcome of the design effort." – Tua Björklund

In this sense, the impacts are broader and can be translated in larger scopes than the organizational one, such as in ecological impacts or social impacts. In order to measure the impacts of a service, a comparison group and enough time passed since the implementation of the service are needed.

2. The subject of evaluation

This leads to another difficulty, which is actually understanding what the subject of the evaluation is. As presented in the literature review, Foglieni, Villari & Maffei (2018) argue that there is a difference between evaluating service design outcomes and evaluating service design practice. This difference was also remarked in the interview, where Foglieni explained how her work deals mainly with service evaluation, or in other words evaluating the service outcomes. Part of her work includes experimentations

around field research, transforming it into an "evaluative research". Foglieni uses the same tools that are used in field research (such as observations, focus groups, questionnaires, etc.) but with evaluation purposes: instead of using these tools just for simply collecting information, they are used to collect data aimed at expressing judgments with respect to what has previously been identified as the key values of the service. In fact, according to Foglieni, Villari & Maffei (2018), the first step of an evaluation strategy would be setting value objectives, thus deciding what is going to be measured. Secondly, indicators need to be created, deciding which would be the metrics suited for measuring the value objectives. Finally, a decision on which tools to use is to be made.

According to Foglieni, the reason why her work concentrates on measuring the results of the service, and not the service design practice itself, is because it would be hard to understand how to measure the capabilities of service design and the effects that these have when introduced in an environment.

"It is less clear how to evaluate the impact of service design, the effect of service design on an organization, a context, a community and so on. It is not clear what to evaluate substantially, and how." – Francesca Foglieni

3. Evaluation is context-dependent

The tight dependence of service design with the context in which it is implemented and the difficulties to isolate the service's impacts from such context are the most difficult barriers to overcome.

"Design isn't really something that you can box neatly in a certain domain. Cross-functional collaborations make it hard to isolate the effects brought about by just design" – Tua Björklund

It's challenging to understand the full picture of design in an organization and it is quite hard to understand if the service results are only tied to the implementation of co-design. How can we tell that co-design methods had an influence in the project outcomes if there are several factors that come into play all together? In the end, it is impossible to attribute a numerical factor to design in the project equation. The more design is embedded in an organization's strategy and ways of working, the harder it will be to isolate it from the rest.

4. Barriers to service design evaluation

According to the interviews, a reason why evaluating design practice is difficult is because designers are very often the minority within the organizations and represent the new ways of working. In this sense it might not be as easy to listen to such a disruptive role inside the organization and it might be hard to track the designers' work. In addition, sometimes service design and co-design are not understood in the organizations, which are skeptical to invest in new approaches if they do not clearly see the benefits.

"If you come up with the idea of measuring outcomes, it doesn't resonate with the people you are trying to communicate it to. So first you have to embed the thinking, what is the purpose and how is co-design actually helping them to achieve their goals. Business managers in all the organizations have a certain goal to reach and they are not interested in anything else than things that are helping to promote that goal." - Taina Mäkijärvi

In addition to this, evaluation is an activity that requires a lot of resources in terms of money allocated, time spent, and people involved.

Even if the organization manages to accept the co-design work and find funding for running the evaluation, often is the case that designers alone do not possess the skills and capabilities necessary to run an evaluation.

This is also true in terms of organizational language, meaning that designers need to speak the language of the organization, allowing the people working there to understand the important takeaways from the evaluation activity.

They key to success is forming a team with different expertises (designers, engineers, managers,...) to locate which metrics are the most

meaningful for the company.
Another difficulty for evaluation stands in the type of data collected by the organizations. Even if some companies are already collecting data, it is not said that it is currently used in a fruitful way.
A crucial step is deciding which value objectives to monitor and select the right tools to collect information. These are difficult tasks, but this is where designers can work together with stakeholders from the company to co-create shared evaluation objectives.

"Defining value objectives is certainly the most difficult part because it is not easy to identify them, and it is not easy to negotiate them with those who commission the project. There are no right or wrong objectives, but there is an objective that must be shared. And then starting from this, once this is clear and well defined it becomes easier to define the indicators and tools that must be used." – Francesca Foglieni

Although there are several barriers to overcome in order to evaluate the impacts of service design and involving users in the development of services, the interviewees expressed great interest in the topic, which seemed of actual importance for them. The urge of further research on the topic is underlined by the several benefits of impact evaluation discussed in the interviews.

5. Reasons for service and co-design design evaluation

One reason to evaluate that was recognized by three out of four interviewees is the need and willingness to get external approval or validation. This happens by utilizing commonly used metrics so that a company can be compared to other organizations. By ranking themselves against each other, companies want to look competitive in their offering. This evaluation goal though, was seen as a marketing purpose, and not closely associated with evaluating service design performance within the organization.
Advocating for co-design practices and for a service design approach internally is another reason why to evaluate the impact of co-design. By showing proof of effectiveness, other departments might be more inclined to try co-development, since it already worked for somebody in the organization.

"In the research we have been doing, the designers that were working in organizations where they are the clear minority, highlight that the best convincer is personal experience . . . . Seeing is believing. Usually people that have seen design methods and collaborated with designers don't need the proof that much, they see the effectiveness and are convinced." – Tua Björklund

"We found that in that case [case study involving an hospital] and in some other case studies with the city of Helsinki, the reason behind if they had an impact was because of an inside agent, an inside person within the organization that really understood the design process and that was in close collaboration with us." – Kirsikka Vaajakallio

Another reason why to evaluate is building factual grounds on which to take decisions. This is especially true for "milestone decisions", for instance where to invest or which concept to choose.
Of course, one reason why to evaluate is to improve, for example improving the design processes within the organization.
This can be achieved by adopting service design tools with an evaluative goal. For example, a journey map becomes not only a tool to see where and what touchpoint to improve, but it is also a means of comparison between the previous user journeys and the current one, becoming a tool for evaluation.
Summing up, improvements can happen at many levels and in many forms. From improving the performance of the design project, to improve the process and the practices within the organization.
Another reason why to evaluate the impact of co-design is to gather proof: from being accountable of one's own choices, expenses, and process to collecting proof for the media.

Currently, the kinds of evidence that the companies are most interested to keep track of and measure are financial measures, customer satisfaction and employee satisfaction. Nevertheless, the interviews uncovered other and different reasons for evaluating co-design and co-design impacts, as seen in this section. The three main reasons are: advocating, learning and improving, and accountability. Understanding the motivations for evaluation is a key element for starting a discussion around the specific metrics and indicators needed for the different purposes.
Figure 11 sums up the findings from the interviews concerning the reasons for evaluating co-design.

5.1.ii Evaluating the Impact of Co-Design

Pondering the reasons behind evaluation, as well as benefits and barriers to its implementation is useful to understand the why. But, in a setting of an organization, who would run the evaluation activities? And is there a right time to start implementing the evaluation framework? These questions were addressed during the interviews, and the answers were very similar.

1.Culture of evaluation

Firstly, no major theoretical differences in carrying out an evaluation addressing the impacts of service design in the private or in the public sector were found. Even though it is common to think about the public sector as more restricted in terms of resources' utilization, this can also reveal true in specific environments of the private sector. Similarly, public sector tends to be more closely monitored from external parties, but again this might be applicable to private organizations as well, for example for the companies who are in the stock market. Apart from practical differences in the amount of resources, public organizations are often very keen in showcasing stakeholders and customers value. Nevertheless, the financial aspect is still important, even if it might not be at the focal point.

Figure 11. Reasons and barriers to co-design evaluation: a visualization.

Generally, the culture of evaluation is more developed in the public sector, as public programs are often already the subject of measurements aimed at impacts' evaluation and improvements. This might come from the fact that the end users in the public sector are closer to the organizations, whereas in the private sector sometimes there could be different layers between the end users and the decision makers or designers, which complicates the measuring activities.

2.Who should run an evaluation?
The agreed answer pointed at a multidisciplinary team. Joining forces from different expertise and departments helps addressing the communication problem explored in the last section, allowing an effective "translation" of knowledge into a language that would speak to the employees as well as the top managers. In addition to that, partnering with different departments inside the organization is helpful to shed light on what are the important indicators to assess in the organization. Focusing on understanding what really matters for the organization saves time and resources, making business units and organizational leaders powerful allies in planning and executing the evaluation. On the other hand, it is crucial also to involve designers because they would be able to spot different kinds of potential impact that could be measured. The ideal setting would potentially include also an evaluation expert in the team.
Currently, new ways of evaluation are being tested and investigated. For example, Hellon developed Aino, an artificial intelligence program which integrates different types of employee experience measurements and other metrics. It can process both enormous quantity of qualitative and quantitative data, allowing the designers to focus on the actual problem-solving activities, and not on compiling huge amount of information. These technologies open up a whole new range of possibilities in impact evaluation that would be worth exploring through further research.

3. Evaluation timing
Evaluation should be a continuous light and systematic activity with more in-depth evaluation moments every now and then. Moreover, different moments in the co-development process require different methods and focuses of evaluation. For example, the evaluation goal will be different when assessing a concept or prototype during the process, rather than evaluating the final service results after the implementation of the service. Similarly, also evaluating the collaborative performance is possible both after the implementation and during the process, as supported by the literature (Voss, 1992; Yin et al., 2011).

5.1.iii Metrics used in Service Design and Co-Design

During the interviews, methods and metrics that could be used in service design are discussed.

1.Evaluation methods
All interviewees agreed that the ideal methodology comprehends both qualitative and quantitative methods. Since evaluation activities require the employment of several resources (money, time, effort), doing pre-evaluation activities using quantitative methods allows the team to individuate the problems on the surface, to later dive deep utilizing qualitative methods. In addition, by using both qualitative and quantitative methods, the picture depicted will be more nuanced and will offer data on the situation, as well as cues for improvement.
Several methods were mentioned in the interviews. From the qualitative side, tools such as recorded presentations, storytelling, narratives, anecdotes, retrospectives, steering meetings, interview studies and self-reported data were cited.
The importance of collecting qualitative data was stressed several times during the interviews and they were always seen as complementary to quantitative tools, the latter allowing for direct and numerical data, the other allowing to uncover motivations and ways to improve.

In addition to examples from qualitative methods, quantitative methods were also referred to during the interviews. Examples are surveys and open excel spreadsheets to gather employees' comments.

2.Metrics and indicators

Besides discussing the methods for collecting data, opportunities for metrics and indicators were also touched upon. The first three interviewees, focused on gathering more information to build background knowledge, stated it was difficult for them to share examples of metrics, without concentrating and knowing a specific case. In fact, they agreed on the fact that the metrics should be crafted ad-hoc every time, starting from what are the important values for the company or organization undergoing the evaluation.

"Beyond the tools that are used, either qualitative or quantitative, the data that is collected must be gathered based on the value objectives that are co-created at the beginning. These indicators must be created every time, they are not given a priori, it's not something that you can find on the internet" – Franscesca Foglieni

Even though metrics can be tailored on a particular company, those are just representative of one category of metrics. Another category comprehends the more common metrics, shared amongst different organizations, that may be on a national or international level. The indicators that are regarded as common practice to measure, such as NPS (Net Promoter Score), Employee Experience surveys, CSAT (Customer Satisfaction), and similar, are helpful in order to rank the companies and compare them. In most big organizations, the data required by these metrics are indeed collected. On the other hand, a company can decide to create specific metrics based on their organization or on internal projects. These tailored metrics can depict better the company's values and can showcase its progress through time. The two categories of metrics have different purposes. Whereas the common metrics are used mainly for external communication and branding purposes, the ad-hoc metrics can be used to foster development and improvement, as well as highlighting the best practices to follow in the organization.

A problem with the metrics already in place in the organizations can be found in the fact that not always the data is collected in the best way because the questions designed might be misleading or asking for the wrong kind of input from the people taking the survey. This highlights the complexity of evaluation: it is not enough to collect data superficially hoping to adapt it into a meaningful metric. The evaluation needs to be designed carefully and a specific evaluation framework should be in place.

On the other hand, a problem with specific metrics is that they are not comparable. For instance, if a company had different metrics for several different projects, it would be impossible for the organization to compare the data collected. The metrics would serve only the specific project they are tailored on and would not be meaningful for any other process. This represents a waste of resources and is not beneficial for the company in the longer run.

Therefore, the approach suggested during the interviews recommends the companies to incorporate both categories of metrics in their evaluation strategy. The companies should continue measuring what they regard as most beneficial for them, meaning they should continue implementing the metrics that are already in place in the organization. In addition, they should consider adding specific metrics tied to the strategy of the company, their overall aims and goals. By crafting new metrics that are tied to the company overall aim or strategy, one can have a compromise between more specific metrics and indicators that are "cyclical" enough to be compared in a few years range.

The implementation of metrics allows interesting result when they are used in comparison with other circumstances. For instance, after creating a set of relevant indicators, the company could compare projects that have used a co-design approach and other projects within the same

organization which decided not to implement such approaches. In this way the organization can understand how co-design is helping (or not) to perform better in regards with a set of specific metrics. Another option could be running a pre-evaluation aimed at gathering evidence on the current situation, in order to compare the current data with the one gathered after the implementation of a co-design intervention.

When coming to metrics in service design, one suggestion was measuring accordingly to the level of maturity of the organization. Considering the Design Ladder, for instance, organizations on the first level, meaning no design is implemented at all, cannot measure service design or design thinking, according to Björklund, since they are not using any design. At that point, it might be interesting for organizations to look at benchmarking in order to form an idea on how they can start implementing design approaches in their own company. On the second stage of the ladder, where design is seen as stage of the ladder, where design is seen as a final polish, measurements that take into consideration the financial value brought by the implementation of service design are at the focus. On the other hand, on the third level, where design is already established, it might not be enough to showcase just financial value, but other kinds of measurements aimed at evaluating customers and employees' satisfaction, as well as the improvement of internal processes, might be needed. Finally, on the fourth stage of the ladder, where design is already embedded on the strategic level, it becomes hard to isolate the effects of service design due to the cross-functional collaborations and the variety of factors that come into play.

"What makes it difficult to measure design specifically is that if you really do have widespread design on a strategic level it's really hard to isolate the specific impact of design." – *Tua Björklund*

Different companies are at different stages of their journeys in the implementation of service design, and this is necessarily translated into different measurements. Interestingly, the reasons why a company employs service and collaborative design approaches might represent the overall goal that needs to be assessed through a series of Key Performance Analysis (KPIs) aimed at that particular objective. For instance, the reason behind implementing collaborative design approaches might be fostering organizational change, moving towards a more collaborative and user-oriented workplace. In order to assess if the transition is happening and if it is successful, the company might decide to have a set of KPIs aimed at monitoring this process. Adding to the discussion, the literature shows a vast array of benefits brought about by co-design processes. Further research on ways to cluster and showcase these benefits and effects in a meaningful and established way could be helpful for organizations. Companies would have the means for realizing they could look at the benefits of co-design as not just potential effects, but goals and aims which can be realized only through a co-design process. In this sense, evaluation would be useful in order to monitor and assess the transformation process of the company.

3.Organizational change actors

Another consideration concerns the organizational change actors. During the interviews, Kirsikka Vaajakallio shared how in her experience, change is easier when there is an "inside agent" in the company, who is closely working with the consultancy (in her case) and understands the motivations and ways of working of co-design and service design. These inside agents are oftentimes the bearer of organizational change and they need to be familiar with co-design and understand the implications and benefits. Inside agents are the internal advocates of co-design, acting as a bridge for co-design knowledge and experience to enter the established ways of doing of the organization, bringing about change. In this sense, evaluation is useful to crystallize knowledge in the other actors who are not inside agents. Collaborative performance measurements can be helpful to contemplate the

the benefits of the process in the service delivery. For instance, during the interviews Taina Mäkijärvi suggests to adapt the retrospective element of Agile project work to include co-design as one of the aspects of reflection, providing an opportunity to evaluate co-design impacts after each sprint, or each project. Furthermore, it is to be considered how in some companies promoting co-design is a challenge itself. Sometimes, co-design is not fully understood by companies and they do not clearly see the value of implementing such approaches in their operations.

"Co-design comes from service design, from the methodology, but then if you think about running an operating company, it is not said that this thinking is embedded in how things are run over there." - Taina Mäkijärvi

On one hand, it is hard to evaluate something that is not valued in the first place, but on the other hand, evaluation might help in raising awareness on the value. Measuring co-design performances might be helpful in order to underline the effects, and importance of such approach, eventually fostering positive responses on the process. On another level, if companies start measuring co-design continuously, more and deeper learnings will be produced, which will be beneficial for the company's improvement.

5.1.iv Conclusions

During the interviews, common metrics such as employees' and customers' satisfaction were briefly discussed, but no metrics aimed at measuring co-design or co-design impacts were encountered, supporting the research gap individuated through the literature review.

Additionally, it is difficult to understand what is going to be the object of the evaluation, also given the fact that isolating co-design activities and impacts from the context and other factors might not be possible. Despite the several barriers, the interviews confirmed the need observed in the literature review of gathering evidence supporting co-design. The evidence required are different for different purposes of co-design, as well as for the different maturity level of the organizations. Therefore, this research argues that more structured evidence and ways of gathering data aimed at co-design and service evaluation are needed. A more organized framework for service and co-design evaluation could be beneficial for organizations, which would have at their disposal a tool that would foster continuous learning and development, as well as helping in keeping track of the process.

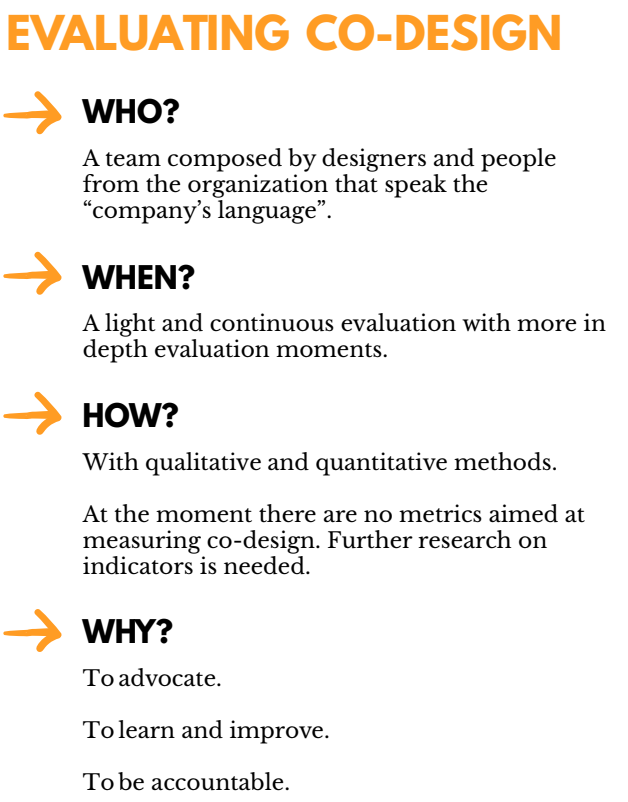


Figure 12. How, Why, When and Who of Co-Design Evaluation according to the four experts´ interviews.

5.2 School Visit´s Analysis

After presenting an overview of the experts´ interviews analysis, this section showcases data and observations from the fieldwork in the school setting.

The fieldwork was carried out during one visit to a basic comprehensive education institute in the Helsinki area. The name of the school and the names of the participants in the research is kept anonymous.

The research was carried out during a Buddyschool class at the beginning of March 2020. Buddyschool classes consist of a gatherings of children from different grades: the younger students are pupils, whereas the older students are tutors who dedicate time to teach the younger students. The specific class that hosted the fieldwork featured exercises such as reading training and math games.

During the class, two researchers were present, observated the class environment and conducted brief interviews with the children and the teacher in the class.

In addition, two more interviews with teachers were carried out at the end of the Buddyschool class in the teachers´ room.

5.2.i Analysis of Observations & Interviews with Children

In total, ten young children from second grade and four older students from sixth grade were present in the classroom, as well as the teacher. The class started by introducing the older tutors to the young pupils, by writing their names on the whiteboard. The Buddyschool coordinator joined for the first part of the school visit but did not stay for the whole time. After the presentations, the children went to eat together in the canteen. When they came back, the actual classroom activities started. The teacher divided the younger children and paired the groups with an older tutor. The pairing was based on the judgement of the teacher, who knows the school progress of the children as well as their personalities. The activities performed in class included reading exercises and Math Bingo. While the preparations and the reading activity were carried out by teacher and children, the researchers limited to observing the environment, whereas during the Math Bingo some brief group interviews were conducted.

The aim of the interviews was exploring the experience that children have of Buddyschool and the learnings that their participation in the program is fostering. For example, two out of the three sixth grade student reported how they did not learn "new things related to school", but they learnt "social skills". On the contrary, the fourth older student reported that Buddyschool is also helpful for her academic skills: since she is not very good at math, she felt that being part of a group with younger students improved her calculus skills. The awareness that the older students had about their participation in Buddyschool and the way they were able to articulate the wide array of different learnings was an unexpected finding. In addition, one sixth grade student shared how, when asked who would like to volunteer for being a Buddyschool tutor, all the children in her class raised their hands. This shows evidence of the excitement that older tutors have for being part of Buddyschool. Nevertheless, the excitement can be underpinned by different motivations to join. For instance, one older tutor shared how she joined Buddyschool because it allows her to skip her classes and it is more fun than following regular lectures. Still, even if motivations for joining might not be driven from willingness to learn and improve, joining the program seems to be a first step towards a path of personal development for the older students. One of them shared how she joined the program because of her great memories of once being a younger pupil in Buddyschool. Generally, the atmosphere in the groups was positive and open, even if the younger students were not all participating at the same level. Some of them were struggling to cope with the

group dynamics and were shyer than others. Interviewing the younger pupils was more difficult than talking with the older students because it was harder to engage with them. The majority did not know what to say when asked about their experience in Buddyschool. In addition, the younger children were very focused on the classroom's activities and did not seem keen on talking with the Finnish researcher.

As for the role of the teacher, we could observe how the initial idea of Buddyschool represented by the sentence "help for homework, no adults allowed" changed and evolved. A teacher was present throughout all the Buddyschool time and she managed and supervised the activities, offering help when needed. Clearly, in the classroom environment the teacher was in charge, as demonstrated by the fact that it was the teacher who presented the older tutors to the younger students: the teacher still holds the authority of introducing the tutors, allowing them to step in the delicate classroom environment and be part of the activities with the younger pupils.

During the activities, two typologies of learning flows were observed. One being the academic learning, mostly happening from the teacher to younger student, as well as from older tutors to younger pupils. Exceptions in this flow can be represented by the academic learning passed from the younger children to the older ones, as reported by the experience of one of the older tutors interviewed. In addition to the academic learning flow, we could also observe a personal learning flow. The main direction of the personal learning flow was from the older students to the younger ones. The tutors are regarded as role models to follow by the young children, and therefore a positive peer pressure feeling is created in the class environment. Moreover, personal learnings were also generated from the teacher to both older and younger students. Personal learnings observed in action were, for example, the ability to listen and be patient, the ability to support others, communication skills being used effectively, and the ability to deal with different situations, understanding how to behave in them (for example being a tutor to a younger pupil who moved to Finland the last summer, in contrast as tutoring a child born in Finland). Although both learnings flows move in different directions and start from different actors, the teacher was clearly still seen as the main authority in charge of the classroom. She is the one managing the division of the groups, deciding who works with whom, she decides on the class activities, and she offers support and indicate the right way of doing an exercise. In this sense, the teacher in the classroom has the power to manipulate and foster the knowledge transfer, both academic and personal ones.

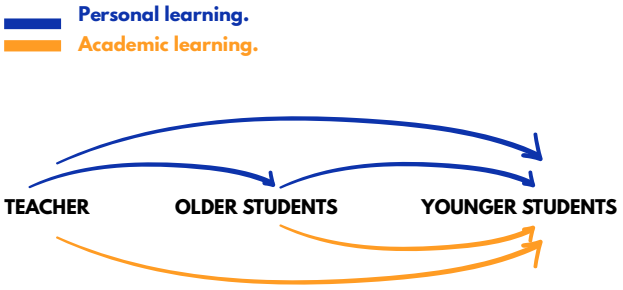


Figure 13. Academic and learning flows observed during the Buddyschool class.

Figure 14, 15 and 16. Pictures taken during Buddyschool class.
cc Maija Astikainen / City of Helsinki

5.2.ii Analysis of Interviews with Teachers

During the school visit, three teachers were interviewed. One of them had just recently started working in the visited school and was interviewed during class time, whereas the other two teachers held a more senior position and were closer to the idea of Buddyschool. During the interview with the new teacher, when asking if she was aware of the process that led to Buddyschool, she answered that she did not know anything about the co-design process leading to Buddyschool and that these ideas were not presented to her in any circumstance, not even during the briefing received in order to become a Buddyschool teacher. Similarly, the two senior teachers could not properly articulate the process that took place before the implementation of Buddyschool. The first part of the interview was conducted in English, and the senior teachers' answer to the question "Were you part of the original process of co-design?" was promptly affirmative. During the interview, some communication problems arose and in order to keep the conversation as clear as possible for the teachers the language of the interview switched to Finnish. During the debrief analysis it emerged how, in reality, the teachers interviewed were not part of the original co-design team, despite having replied yes multiple times.

This highlighted how the perception of the project's beginning according to the two teachers was different than the actual starting of the project. In their own view, the project started when Buddyschool started to be implemented in their school. Even though they recognized that another part of the process happened before the implementation, it seemed they could not link the two parts together to create a continuous timeline of the process.
The findings from the teachers' interviews are further articulated and combined with highlights from the literature review in the next chapter, *6. Research Conclusions and Proposal*, which provides an overview of a critical comparison between fieldwork, literature review, and experts' interviews.

5.3 Migrant Youth Helsinki Interviews´ Analysis

The method of semi-structured interviews was also used as a means of preliminary exploration of the case study before conducting the school fieldwork. The interview series with Migrant Youth Helsinki comprehend three formal interviews featuring the service designer who was part of the co-design process, the project manager of Migrant Youth Helsinki, and the product owner of Buddyschool. In addition to the more formal interviews, part of the reflections highlighted in this section also comes from the meetings and several email exchanges that were not a structured part of the methods.
In the next sections, the main findings emerged are presented.

5.3.i Reasons for Success

The interviews with the staff highlighted the reasons for success of Buddyschool. The main motivation that was recognized by the three interviewees is the fruitful collaboration between the team members. In fact, they reported being a strong team from the very first phases of the project. Moreover, the implementation of a team composed by a service designer and the project manager and product owner was seen as an innovative way of working. Usually, municipalities utilize co-design and service design through external consulting work, but in this case a team composed by a lead service designer and the project manager was formed right from the beginning of the project. In this way the project manager could be part of the whole process, allowing her to understand the essence of the work and the structural decisions.
During the interview with the project manager, she highlighted how this project represented a great learning experience for her and how they

implemented a new way of working that was rarely implemented before in City of Helsinki.
Another reason for success that emerged from the interviews are service experiments, or service prototypes. Testing the execution of the idea in a real-life scenario was deemed an essential step in the process. This allowed to iterate the idea and to understand how the service would fit the experiences and the dynamics of the people using it in a real context. The culture of experiments is not new to City of Helsinki, and this case further proved the usefulness of testing a concept in a real life scenario.
Thirdly, a motivation that led to success that was highlighted several times in the interviews is the fact that the idea for Buddyschool came from teachers. This means that the idea is not imposed from outside, but on the contrary the idea is generated from everyday users of the service. Showing ownership of the idea helps in seeing the benefits and being more motivated in implementing it. Collaborative idea generation will be discussed further on in this section as part of the value of co-design during the design phase.
Similar to the idea generated from "insiders", another reason for success was the network of teachers that started forming in the implementation phase. During the interview with the lead service designer, he reported how teachers were helping each other during the initial phases of the implementation, by sharing advices and becoming "trainers" and helping teachers that were new to the service. He also underlined how, even if the role of the project owner is crucial to get the service started in a new school, the training could only be fruitful if performed by other teachers. As Buddyschool pupils learn in a more engaging and rewarding way if accompanied by peers, similarly teachers can relate more to other teachers. This seems to indicate the presence of trainers amongst the group of teachers. In addition to the role of the trainer, a slightly different but similar role can be found, the ambassador. Both trainers and ambassadors share a willingness to promote Buddyschool, but the role of the trainer is focused on promoting the best practices, the set up and the right utilization of the service, whereas the ambassadors also promote the ways in which Buddyschool was created, advocating for the whole process. A more in-depth presentation of the teachers' role in Buddyschool is presented in the next chapter, 6. Research Conclusions and Proposal.
Finally, the interviews underlined the importance of a fruitful goal setting. Throughout the process, values objectives were co-created amongst the Migrant Youth Helsinki team and the co-designers coming from the teaching body and the migrant youth. The reason for success was dividing the bigger objective in smaller steps. For instance, the project was not approached by just focusing on the bigger scope which is limiting the school dropouts after 9th grade, but the goal setting was divided in smaller phases. During the interviews it was highlighted how it was important that these phases were ambitious enough to allow the project to push the boundaries, achieving big goals.
It is interesting to highlight how the reason for Buddyschool´s success encompass the elements of co-design as described in the second chapter, *2. Research Setting*. The pillars of co-design adapted from the NSW Council of Social Service (2017) presented before underline how a co-design process is inclusive, iterative, participative, solution focused, and enabled by co-design tools. The interviews with Migrant Youth Helsinki staff stressed how these characteristics that make Buddyschool a co-design project, are also the strong points of the whole process.

5.3.ii Value of Co-Design

In general, the staff from Migrant Youth Helsinki all seemed supportive and aware of the co-design process. They talked about it with high regards and they highlighted how the process represented a learning experience for them. The interviews highlighted how the staff from Migrant Youth Helsinki could see the value of using co-design. Specifically, two different stag-

es with different values of co-design emerged: the value of co-design in the design phase and the value of co-design in the execution phase.

The design phase is comprehensive of the steps taken before the buy in from the executive level. The first step is represented by the identification of needs and problems, which leads into the idea generation. The design phase also includes the validation of the problem-solution fit through experiments and service prototypes. During this stage, co-design is mostly seen as co-development of ideas and concepts, as well as participating in collaborative validation activities. This includes brainstorming together, using co-design workshops and interviews, setting up experiments in the school environment and asking for constant feedback. The value of co-design in this phase that emerged in the interviews is that it allows the staff to open their eyes on different possibilities and get in touch with different realities. They can be aware and participate in the everyday problems of different stakeholders and users in the process. In this sense, it was recognized how engaging users allowed Migrant Youth Helsinki staff to see the reality of the problem and frame it better, leading to a solution that fits the needs of the user in a more effective way. In addition, the value of co-design in the design phase is also located in the collaborative idea generation, that allows the users to feel engaged and active participants in the process. Moreover, if the ideas are generated within the co-design setting, it means that the users have ownership of the ideas, which will be better accepted rather than the ones coming from outside, imposed from the school system. Interesting further research could possibly deal with comparing Buddyschool, which was generated from a co-design process, with another program implemented in the schools were teachers and school workers did not have a chance to express their opinions about. It would be interesting to compare how the two services are doing, if one is accepted and is performing better than the other, or not.

> *"A lot of concepts are pushed into the school system without understanding the daily structures and dynamics in the school's work. She [a teacher participating in a workshop] told me that usually she feels bad because she is seen as the obstacle in making the school a better place. In fact, usually teachers are being told what to do and how to act but they sometimes don't understand what the reasons behind some particular decisions are. This is why the program is designed on a paper on a very specific level and the teachers can't relate to it".*
> *– Iika Lovio, Service Designer*

The second phase is represented by the execution phase, which happens after the buy-in from the executive level. This stage includes the goal

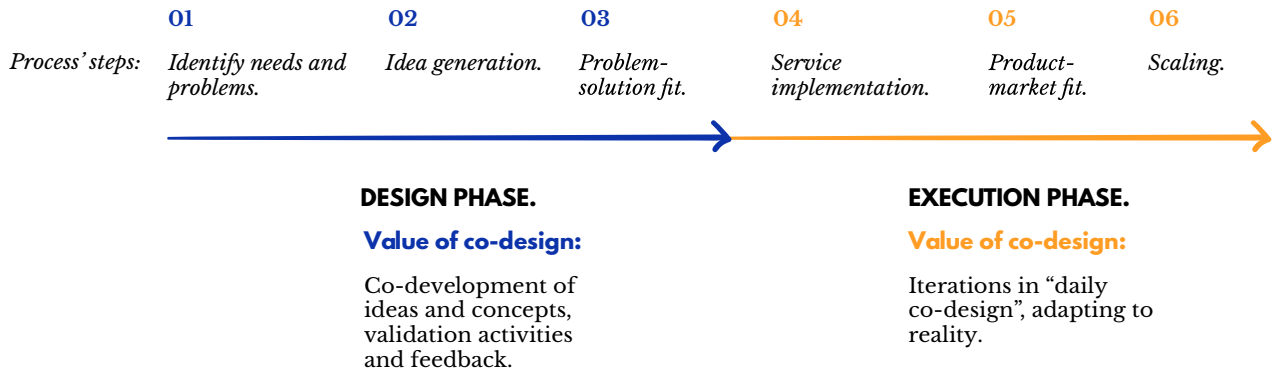


Figure 17. Co-design value in the Buddyschool process. Visualization based on interviews with migrant Youth Helsinki´s Service Designer.

and KPIs setting for monitoring the implementation of the service and aiming at scaling it, as well as iterating the way the service is delivered according to the different environment it is placed in. According to the interviews, in this phase co-design and goal setting should continuously support one another. The goal setting should be ambitious so that the ideas that would not spur the project to the goal can be eliminated. For instance, when planning the scaling up of Buddyschool, the numbers of schools which the service should aim to cover in the next years was set very high. In this way, the team was motivated to reach the goal. Moreover, the value of co-design in the execution phase can be found in its iterative nature. During the interview with the service designer, co-design in the execution phased was referred as "daily co-design". This term underlines the intuitive changes made in some features of the service in order for it to fit better in the school dynamics and environment. This means that when the service is implemented and finally meets the real world, the co-design activities do not stop, they just translate on a different layer, becoming embedded in the delivery of the service. Therefore, constantly adapting to the reality by iterating the service characteristics and delivery is a form of co-design, even if the people who create these changes might not be using the word "co-design".

5.3.iii Evaluation in Buddyschool

During the process there was no formal framework for evaluation. Nevertheless, feedback was constantly gathered and acted upon. Overall, the reasons for gathering feedback were testing ideas and service prototypes and collect data on the service performance after implementation.
The first plan to follow up on the bigger scope question on how Buddyschool would impact the lives of students was to track students' social security numbers. The original service idea planned for only students from 9th grade to be tutors, and students from 1st or 2nd grade to be pupils. In this way it would be easy to track their social security number and to implement strict test conditions, where only defined target groups could participate to maintain the comparison valid. This plan proved soon to be not possible to realize. Firstly, it would have meant restricting the pool of candidates for teachers to choose the children from, and secondly it would have been difficult to isolate the influence of Buddyschool on the lives of children. For instance, children's life might have taken a positive turn because their unemployed parent found a job. Therefore, the plan of tracking the social security number was never implemented and the staff preferred concentrating on other methods, such as interview studies and questionnaires. In general, the feedback gathered was aimed at testing the service features during the design process and at assessing the service outcomes after the implementation. No feedback was gathered on the design process. The assessments were mainly produced in the form of narrative stories and data showing the progress of the service delivery and students. Examples of data collected are the increasing number of schools enrolling in Buddyschool, or the improved grades of the children. Most of the energy, though, was concentrated on gathering narrative feedback, as it is the format that allows for more nuanced explanations and the understanding of users' reasons. The interviews were carried out by the product owner in order to collect personal and narrative stories and anecdotes. The interviews were used to collect quotes and data to be used to further improve Buddyschool, or to showcase during update and steering meetings.
After the implementation of Buddyschool, a survey was delivered to teachers in order to monitor how the service was performing. The aim of the survey was researching improvements in the academic and personal skills of he children, as well identifying improvements in the school atmosphere and effects on the teachers' body. The questionnaire was answered by 40 teachers and it allowed to rate a statement from 0 to 5. The teachers were asked to give their opinions

on the effects of Buddyschool on young pupils, on older tutors, and on the school community. Overall, the results show how young people have taken on a more responsible role, showcasing improvements in learning and motivation. In addition, the survey highlighted a close interaction and collaboration with students and teachers, which promotes a healthy school community. In addition, the survey underline how the older tutors show responsability towards their role and the younger pupils.

Although the questionnaire collects useful data to reflect upon, it is important to note how there might be different interpretations of the scale (0 to 5) depending on the teachers. It would be interesting to explore the idea of a questionnaire with more fixed parameters, where the indicators provide objective statements for every level of agreement. For instance, the statement "Working in Buddyschool has improved peer mentors' school motivation" would need to feature the characteristics of an improved school motivation and the behaviors of a student with higher motivation in coming to school. This would be helpful to craft a scale that would include relevant details on the behavior of the children. In this way the results would still be based on the observations and opinions of the teachers, but they would be provided with more guidance and help in assessing each statement.

In addition to interviews and surveys, the positive welcome and the excitement of teachers and children are also evidence supporting the success of Buddyschool. For instance, the teachers tried out the solution as soon as the idea of Buddyschool was proposed during the workshops. An example from the reaction of the children is that they were spreading the word around school and everybody wanted to participate in Buddyschool. There were also episodes of sick kids coming to school anyway because they did not want to miss Buddyschool.

During the process, several Key Performance Indicators (KPIs) were taken into consideration, but not tracked. The only indicator used to track the collaborative performance was the number of people attending the workshops on a continuous basis. During the initial phases of the implementation of the service, the numbers of children coming back after participating in one class was monitored. In order to test the viability of the service, KPIs focusing on the experience and improvements of the children were taken into consideration. For example, the staff asked questions such as "Do the children feel that the school is more fun?", "Do the children engage with Buddyschool continuously?", "Are the children coping better with school?", "Have the mathematics skills of children improved?". In this sense, the indicators used in the latest stages shifted from being leading indicators to being lagging indicators. Lagging indicators are the ones that might confirm a pattern that is in progress. They are output-oriented and therefore are used to prove the success of Buddyschool, for instance by asking about the performance of the students. On the other hand, leading indicators represent and assess the performance to try and predict a future success. During the interview it was highlighted how the majority of the indicators used were lagging ones (outcome-focused).

Before starting the project, the team conducted an extensive background research on how education affects someone's life chances. The idea of providing a way for coping better with school is based on the evidence that this will indeed help in continuing studies, getting a job and eventually have access to better possibilities than before. Since the connection between enrolling in further education and better life chances was already proven, the staff from Migrant Youth Helsinki felt it was not their duty to reinforce this proven causality through Buddyschool. Therefore, Buddyschool's aim is not to confirm that children who continue studying have better opportunities, but it is based on this demonstrated theory. In this sense, the evidence that must be gathered is not aimed at proving that the children will have better opportunities if they continue their education path,

because that is part of the research on which Buddyschool is built. On the contrary, an opportunity can be found in focusing evidence gathering specifically on Buddyschool results and ways of working to investigate the viability of the service and the reasons for success, including the relationship between the process and the outcomes.

One of the reasons why evaluation was not implemented as a structured activity in Migrant Youth Helsinki is because of the small team. With just a team of three people, dedicating time to measure and evaluate would have meant taking it away from other activities. This confirms the theme of scarce resources and the difficulty of allocating time to such activities found through the experts' interviews. Although the process was not formally assessed, there is a shared feeling of success between the people from Migrant Youth Helsinki who took part in the process.

Indeed, this feeling of success was confirmed from the interviews with the teachers who were all happy with the service. Although the teachers interviewed were able to clearly see benefits in the implementation of Buddyschool, they were not able to link the effects of the service delivery to the co-design process. In other words, they were not able to articulate the important role that the co-design process played in the development of Buddyschool. On the contrary, the staff from Migrant Youth Helsinki fully understood the fruitful causal connection between the process and the outcome.

"We have seen obviously that the process in itself was so valuable that without the process this would not be the end result" – Irma Sippola, Project Manager.

5.3.iv Conclusions

Overall, the research highlights how Buddyschool is a successful service, based on the feedback from teachers, children and Helisnki Migrant Youth staff. The success is brought about by improvements in the children (both academic and personal), as well as a higher service acceptance by the teachers, who feel ownership for the idea. As underlined previously, the reasons for success are also similar to the elements that make Buddyschool a co-design project: an iterative, inclusive, participative process.

No formal or systematic set of indicators and measurements were implemented during the process due to the lack of human resources, but surveys and users' stories were collected as a means for feedback both during the co-design process and after the service implementation. The research shows how evaluation in Buddyschool focused mainly on assessing the service, meaning investigating the results in the execution phase, and evaluating possibilities through different concepts and experiments during the design phase. Nevertheless, the interviews also highlighted how co-design value can be articulated in both phases. The next chapter will explore the

possibility of evaluating co-design in Buddyschool, and not only evaluating the service.

This chapter provided an overview of the findings in the research work with the case study and the experts' interview. In the next chapter, the findings are combined with the background knowledge from the literature review to form insights aimed at finding opportunities for developing a design proposal.

Chapter 06

Research Conclusions & Proposal

6.1 Conclusions from Literature Review and Fieldwork

Whereas the previous chapter, *5. Analysis and Findings*, focused more on showcasing the research findings, the aim of the current chapter is to investigate and identify a solution space, moving towards a design proposal by combining together the data gathered through the case study fieldwork, the literature review and the experts´ interviews.

6.1.i Insights from Buddyschool

The main starting point in compare and combine the case study and the literature review is the fieldwork conducted in the school. The case study offers the chance to investigate the consequences that co-design had on a real service and its development process in order to investigate opportunities for co-design evaluation. In this sense Buddyschool is used as a data collection scenario in order to collect findings that can be used to further craft the design proposal.

Overall, the most critical research finding from the school's fieldwork can be identified in a misunderstanding with the two senior teachers regarding the co-design process. In fact, they were not aware of it, meaning that the knowledge about the co-design process was not passed on from the co-design setting to the real-life environment, in other words from the teachers who were co-designers to the other colleagues.
The actors that participate in the co-design process, who are physically sitting at the workshop table, hold experiential knowledge of the co-design process because they participate in first person.
In this sense, the value of the process for them is articulated in an explicit way and they should be able to see the benefits that the process can realize in the final outcome, as they are contributing personally to the idea generation and feedback process. Nevertheless, the users that are part of the co-design setting are just representatives of a bigger variety of users who will eventually start using the co-designed service in their everyday life. Such actors of the everyday setting do not hold any experiential knowledge of the process, but they can gain narrative knowledge of it. This can happen through a translation process: the experiential knowledge and the process' awareness hold by co-design participants can be shaped into a narrative form and passed on to the actors operating in the everyday setting.
When this knowledge and awareness translation fails to happen, the users that were not part of the co-design setting seem to struggle to articulate the value of the process and its effects on the service delivery. During the interviews, it emerged how in this case the knowledge and awareness translation process failed to be successful. The actors in the everyday implementation of the service were not able to articulate the effects of the co-design process into the service delivery of Buddyschool and consequently felt distant to the inquiries about evaluating such effects.
The concept of knowledge transfer in the co-design process represented here as the shift from experiential to narrative knowledge can be reconducted to a similar view by Suominen & Pöyry-Lassila (2013).
They suggest three different modalities through which knowledge can be shared, arguing for the central role played by "visual boundary objects (e.g. process models, scenario visualizations, intervention session recaps)" (Suominen & Pöyry-Lassila, 2013, p. 10). According to their study, these visualizations used and produced during the co-design activities can be used to foster both monological and trialogical knowledge transfer.
In the first case (monological knowledge transfer), actors present in the co-design setting share the knowledge acquired through their experience to other people inside the organization by using the visualizations as starting point for discussion. In one case study analyzed by the authors, they noticed how these visualizations

can also be tweaked in order to spark further change inside the organization, promoting a trialogical knowledge transfer. This means that service development and co-creation practices are introduced in the organization by the actors participating in the co-design setting. Furthermore, knowledge transfer and sharing foster collaboration on a bigger layer through dialogical effects (Suominen & Pöyry-Lassila, 2013). This highlights how participants of the co-design process can act as ambassadors or champions of the process, translating their experiential knowledge and awareness into a narrative one for the other employees working in the organization. The critical role of ambassadors as change agents inside an organization is also highlighted in the experts' interviews.

The presence of teachers willing to take on the role of ambassadors is also recognized in Buddyschool's co-design process. During the interview with the service designer, he underlined how workshops' participants were enthusiastic about having their ideas heard and develop a service that is not imposed "from the outside". Nevertheless, this comment came from the teachers participating in the co-development process and not from the teachers living the Buddyschool service in their everyday lives.

In fact, the interviews with the three teachers who were not part of the co-development process show how the knowledge transfer between participans in the co-design process and the other teachers was not completely successful.

Even though it is possible to gather evidence about ambassadors, these figures were present early on in the process and the research did not have a chance to include them in the interviews. It would be interesting to track the initial participants of the workshops to further research their current role in Buddyschool and see if the role of ambassadors changed after the service implementation. Unfortunately, it was not possible to retrieve any teachers who participated in the co-design workshop due to time limitations and schools closing for the spread of COVID-19.

The considerations originated from the school visit are hereafter articulated into five takeaways, which represent important data that informed the design proposal.

1. Teachers' categories in Buddyschool process.

Throughout the Buddyschool process, three categories of teachers can be identified. Firstly, the Teachers Co-Designers are the teachers who participated in the co-design process. Teachers Co-Designers were representatives of the ideas and interests of a broader category, the Everyday Teachers. These are the teachers that did not participate in the co-design process but are in contact on an everyday basis with Buddyschool, after its implementation. Finally, Ambassadors are the ones with the willingness to pass

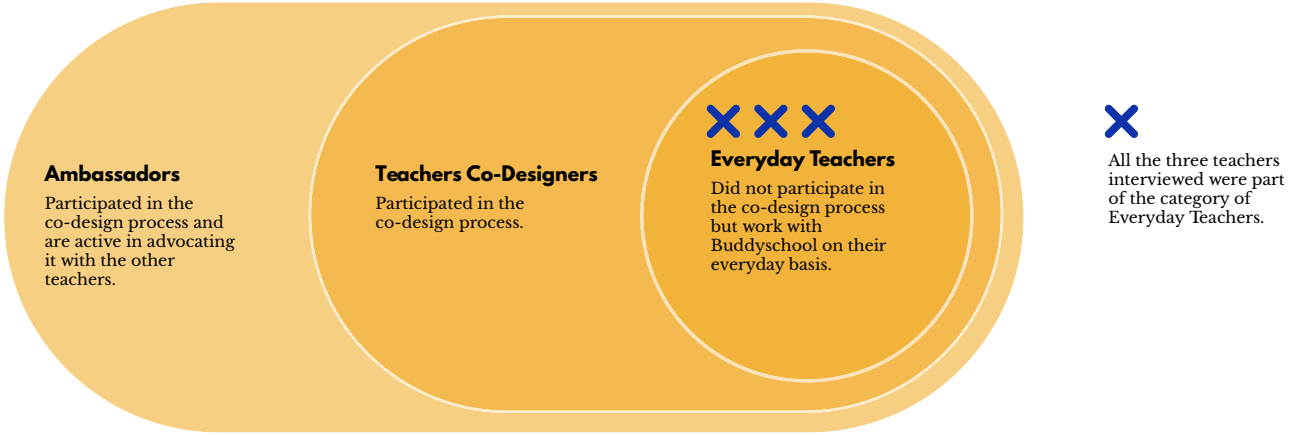


Figure 18. Visualization of teachers´ categories in Buddyschool process.

on the knowledge and the values they experienced during the process. Not all Teachers Co-Designers become Ambassadors, but it depends on personal and contextual characteristics. All Ambassadors are Teacher Co-designers because they need to have experienced the process in order to advocate for it. In addition, all Teachers Co-Designers are Everyday Teachers because after their participation in the process they turned into active service actors, teaching Buddyschool classes in their schools. Given that the Everyday Teachers interviewed during the fieldwork were not fully aware of the co-development process underwent by Buddyschool, resulting in an inability to articulate the value of such process, it can be assumed that the knowledge and awareness translation process didn't successfully take place for the interviewed teachers, thus the formulation of the second insight.

2. Experiential knowledge vs narrative knowledge.

The Teachers Co-Designers hold experiential knowledge and should be aware of the value of the co-design process as they experienced it in the first place. On the other hand, Everyday Teachers did not participate in the co-development, therefore they do not hold any experiential knowledge or first-person awareness of the process. Everyday Teachers can hold narrative knowledge and awareness of the co-design process, as long as a translation from experiential to narrative knowledge happens, as also demonstrated by Suominen & Pöyry-Lassila (2013). A way this translation process can happen is if facilitated by the Ambassadors. Ambassadors hold the power and opportunity of translating their own experiential knowledge into a narrative one, raising awareness on Buddyschool co-design process amongst Everyday Teachers. If this translation process fails to happen, the value of the co-design process is not made explicit for the Everyday Teachers, resulting in a difficulty for them to articulate the process' value and the effects that the process has on the outcome.

The failed translation of the knowledge and process values manifested in the interviewed teachers, as they were struggling to answer questions requiring them to connect the process and the outcome. This means that the teachers interviewed did not have neither experiential nor narrative knowledge and awareness of the value of co-design process in Buddyschool. As a result, they could clearly see the benefits of Buddyschool implementation, but they could not articulate them in contrast to the co-development process. In other words, they could just articulate and see value in the measurements of the service, but not in the assessment of the co-design process. Moreover, their attitude towards the co-design process in Buddyschool was in open contrast with the one demonstrated by the staff from Migrant Youth Helsinki, as well as the teachers Ambassadors. This underlines how the benefits of the process are noticed only by the actors who hold explicit value of the co-design process.

During field research, several benefits brought about by the co-development process in Buddyschool were highlighted, but all of them were touched upon during interviews with the service designer, the project manager or the project owner. On the contrary, the Everyday Teachers interviewed were able to see benefits in the performance of the service, but they were not able to link these effects to the process with a causal relationship.

During the analysis, this sparked further reflection on the way the questions were asked during the interviews and the assumptions that were underling them. The personal and valuable reflection that was generated from this experience of interviewing teachers concerns the assumption that in a context where co-design is widely developed and used by practitioners and researchers, as in the Finnish context, it must be spread also outside of academia and work settings. After talking with the project manager and the service designer working in the project, the expectation was that everybody involved in the project, even frontline workers such as the teachers, would have the same level of

knowledge on co-design and the process. This assumption proved to be wrong, but it also represented a precious learning, as it helped to re-contextualize the research work.

3. Co-design benefits.

Co-design benefits are not seen by who does not hold explicit value of co-design, which comes from holding either experiential or narrative knowledge and awareness of the process. If the translation from experiential to narrative knowledge and awareness does not happen, the teachers are not able to see explicit value in the process, meaning that they cannot articulate how the benefits in the outcome are linked to the process. Therefore, what makes Buddyschool different (the co-design process) is not fully acknowledged.

One counterargument could be that Everyday Teachers do not need to be aware of the co-design process in which they did not participate, as their role is only focused after the implementation phase. Nevertheless, Steen (2013) writes "I advocate organizing co-design according . . . to its ethos, which often remains implicit. We can do so by making these ethics more explicit and promoting reflexivity. . . . By becoming more aware of their environment, participants can organize their co-design more effectively, so that they can jointly learn and jointly create, address problems in the real world, and develop solutions that work" (Steen, 2013, p. 28).

Building on Steen's work, this research argues that the motivation behind making the value of the process explicit not only for the Teachers Co-Designers, but also for the Everyday Teachers means allowing them to be more aware of the role they play in the bigger picture. The teachers interviewed resulted very outcome-focused, and therefore concentrated on the benefits of the service implementation, leaving aside their connection to the process. By being more aware of the co-design process, they could be more aware of their share of ownership of the service and the modalities in which they can continue the co-development even in their everyday work. When interviewing the service designer, he reported that one factor for the success of Buddyschool was that the idea came from inside the teaching body, whereas usually services are imposed from outside without consulting the teachers. But in this case, since Buddyschool is coming from a collaborative and iterative process, the Everyday Teachers could also be owners of the process, even if they did not participate in the initial co-design setting. They can still iterate the service in their own schools, continuing the legacy of collaborative development in their own small co-design setting, which is the classroom. But, if Everyday Teachers do not understand the value of the co-design process

Figure 19. A visualization of experiential knowledge vs. narrative knowledge.

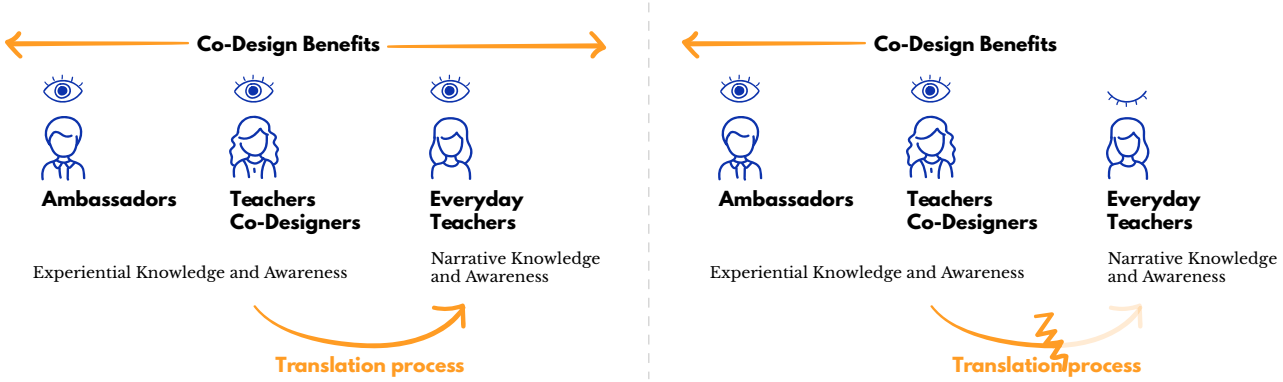


Figure 20. A visualization of co-design benefits´ awareness.

because it is not made explicit for them, they cannot reach this feeling of ownership in order to further continue the development of the service in their own areas of responsibility.

4. Making the value of co-design explicit – "Everyday co-design".

Making the value of co-design explicit for Everyday Teachers and allowing them to see the benefits of the co-design process might be helpful to contextualize their role and opportunities in the implementation of the everyday outcome. By raising awareness on the different means and styles of ownership of the service idea and by broadening the views of the Everyday Teachers, moving from an outcome-focused mindset to a view comprehensive of the process, the boundaries of co-design are pushed out of the co-design setting, fostering everyday participation, input, iteration and improvement of the service. This will allow Everyday Teachers to take ownership in the service development after implementation, continuing the development in their own co-design setting: the classroom.

In fact, during the school visit, it was interesting to observe how the teacher that was not aware of the co-design process in Buddyschool was working and acting in a very similar environment, with very similar dynamics and structures as the setting of a co-design activity. Buddyschool fosters the co-creation of knowledge and the dissemination of learnings on different levels, highlighting how the classroom could be compared to a micro co-design environment. In this sense, the insights generated on the need for teachers to realize their ownership to the service after the implementations seems to find further validation by this observation. If the value of the co-design process in Buddyschool would be made explicit for the Everyday Teachers as well, they would have the opportunity to take ownership of the service and realize the similarities between the process that led to the creation of Buddyschool and the everyday process they undergo during the class activities. Making the process more explicit for the Everyday Teachers means making them aware of the connections between the service results and co-creative nature of the service development. In this way they could be more aware on their personal possibility of further developing the service, acting in their own area of responsibility. Therefore, the fifth and last key takeaway is:

5. Classroom as an "extended" co-design setting.

The classroom environment was representative of a micro co-design setting because of the joint creation of knowledge and sharing of learnings, supported by physical materials and tools. In addition, the roles of the participants are negotiated to allow a process of "joint inquiry" (Steen, 2013, p.24).

Figure 21. A visualization of the two different states of Everyday Teacherss: aware of the co-design benefits, or not aware.

Steen (2013) quotes the ideas of John Dewey, American philosopher, psychologist, and educational reformer, and applies them to co-design: "We can understand co-design as a process of joint inquiry and imagination – as a 'reflective activity in which existing tools and materials (both of which may be either tangible or conceptual) are brought together in novel and creative arrangements in order to produce something new' (Hickman, 1998, p. 169)" (Steen, 2013, p. 24). Similarly, group activities and collaborative exercises in the classroom were aimed at creating new knowledge and transfer it from one actor to the other. In this sense, a process of "joint inquiry" (Steen, 2013, p. 24) could be observed in the classroom. This process was sparked by the tasks assigned by the teacher, which sometimes were articulated in "joint imaginations" (Steen, 2013, p. 24) as the participants of the groups re-arranged and interpreted the tasks according to their own experience. This communal exploration facilitated by the utilization of "existing tools and materials" (Hickman, 1998, p. 169), helped by the facilitation skills of the teacher, produced new collaborative knowledge both on the academic side and on the personal side. Moreover, co-design "is a process in which participants are able to express and share their experiences, to discuss and negotiate their roles and interests, and to jointly bring about positive change" (Steen, 2013, p. 28). Likewise, children were required to negotiate their roles in the group setting, by affirming themselves as tutors or pupils. This happened, in some groups more explicitly than in others before starting the in-class activities, as older tutors employed their communication and facilitation skills by asking the younger pupils if they had any questions or encouraging to start the activities. Moreover, also the teacher was asked to constantly re-negotiate her role in the classroom, switching from the authority figure when briefing the activities, to a more supportive and approachable role when giving advice and feedback on the ways the activities were carried out. Finally, positive change was brought about both on an immediate level as well as a more long-term level, as demonstrated by the interviews with the older tutors. The immediate positive change was observed in the excitement for the activities demonstrated by the younger pupils, who were really enjoying the Math Bingo. On the other hand, the longer-term positive change is represented by the fostered academic learnings and the improvement on the social skills level. In addition, the management of the classroom activities, as well as the time keeping, resembled very much the organizational skills required during a co-design activity.

Summing up, the similarities between a co-design setting and the classroom environment observed during the fieldowork are the following, based on the definition by Steen (2013).

- New knowledge was created and transferred from one actor to the other (from children to children, from the teacher to the children) through a process of "joint inquiry".
- Existing tools and materials were used to produce knowledge and foster learnings on both a personal and academic level (Math Bingo´s material, and books).
- There was a continuous negotiation of roles by both the children and the teacher.
- Positive change was fostered on both a short term (immediate wellbeing and excitement of the children) and a longer term level (academic and social learnings).
- The management activities performed by the teacher were similar to the ones performed by a facilitator in a co-design setting (keeping time, facilitating the classroom activities, assigning tasks, helping children in performing instructions, etc.).

Combining the fieldwork in the case study with the background research of literature review and experts' interviews represents the starting point for reflecting on possibilities of co-design evaluation in Buddyschool. As showcased in this section, the concepts expressed in the case study's insights are also present in the literature. In this sense, the literature and the Buddyschool research have been continuously shaping one another, validating the correspondent findings in both research streams. In general, the case study research highlights how evaluation discourses within Buddyschool have been focused on assessing the service results only. Following the work of Foglieni, Villari & Maffei (2018), it could be argued that in Buddyschool the service was measured (to some extent), but there was no attempt in assessing co-design or the collaborative performance. The lack of research on measuring co-design and its impacts on the service result is also stressed by the literature review, as presented in chapter *4. Literature Review*. Similarly, also the experts' interviews stress the need for further investigating indicators aimed at evaluating collaborative practices. Therefore, there is a common gap individuated in all three research phases (literature review, experts' interview and case study), which is the need for further investigation on indicators aimed at measuring co-design performance and its effects on service development projects.
In order to address this gap through the design proposal, the literature review becomes a critical part in this master's thesis research. The next sections present an overview of the thinking process leading to the design proposal, as well as conveying the data that supports the decisions taken.

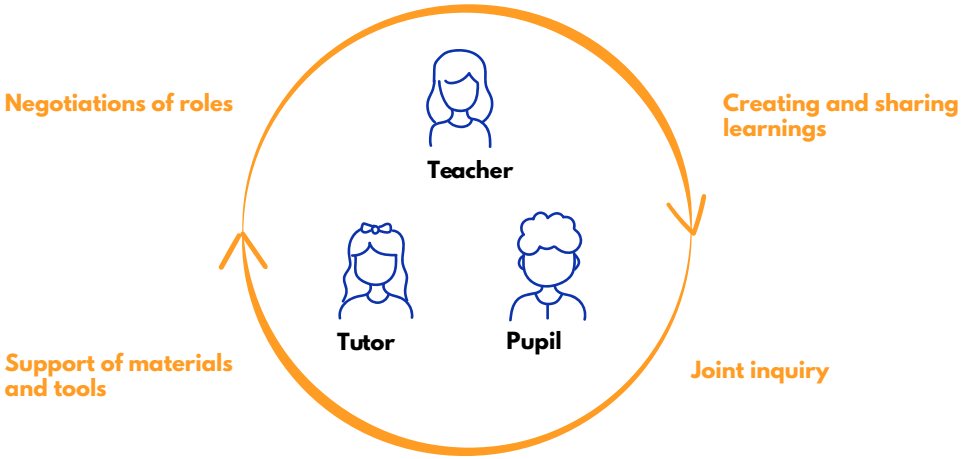
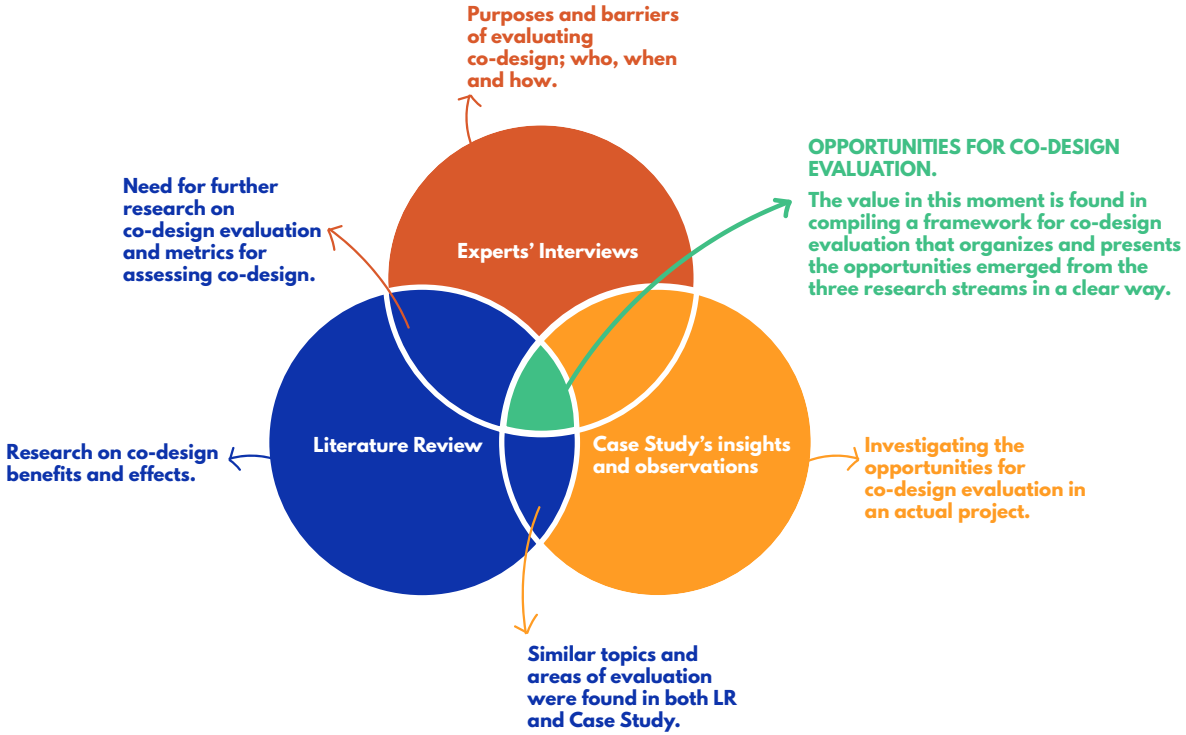


Figure 22. A visualization of the two different states of Everyday Teacherss: aware of the co-design benefits, or not aware.

6.1.ii Towards a Proposal

The research underpinning this master's thesis is composed by three main streams: literature review, experts' interviews, and cases study fieldwork. During the previous chapters, findings concerning the three streams were presented. Firstly, one of the focus of the literature review is the research on co-design benefits and effects, to understand the impact that a co-design approach might be taking into a service design project. Secondly, the experts' interviews articulated the purposes and barriers of evaluating co-design, as well as potential modalities and actors to involve in order to perform the evaluation. Both streams highlighted the need for further research on co-design evaluation and metrics for assessing co-design. On the other hand, the case study fieldwork offered the occasion to investigate the opportunities for co-design evaluation in a real-life service, as demonstrated by the insights which underlined how the topics and areas of evaluation highlighted in the literature review were relevant for the case study as well. In general, the three research streams highlight a common gap, which is the need for further research on co-design evaluation. Nevertheless, combining all three the research streams reveals how the opportunities for co-design evaluation are already out there. Currently, the value can be found in organizing the data that previous research (and the current research featured in this master's thesis) already features.

Figure 23. The three research streams (literature review, case study fieldwork and experts´ interviews) are combined in order to investigate opportunities for co-design evaluation.



Organizing and presenting the opportunities emerged from the three research streams in a clear and straightforward way is seen as valuable for two main reasons. Firstly, it will provide an overview of the state of the art, creating a framework that is easily adaptable to different situations and simple to add to. Secondly, it might foster more interest in the topic and therefore generate further research on co-design evaluation. Therefore, a co-design evaluation framework is chosen as the design proposal of this master's thesis, supported by the evidence of the research. The next section will explain in more detail the thinking process behind the creation of the framework and the theoretical foundations on which it is based on.

6.1.iii Towards a Co-Design Evaluation Framework

The work of Foglieni, Villari & Maffei (2018) presented in the literature review highlighted how evaluation within service design practice can be divided into two broader categories: service evaluation and service design evaluation. These two categories differ in the focus of the evaluation activities. Similarly, this thesis argues that co-design evaluation can also be differentiated in an analogous way, as this research focuses on co-design as a part of the process in service development projects.

Concerning the division between service evaluation and service design evaluation, Foglieni, Villari & Maffei (2018) show how the focus of evaluating a service is found in assessing the outcomes and the results of the process, whereas when evaluating service design the goal is assessing the service design capabilities and skills, as well as evaluating the service design process. In a similar way, evaluating co-design can also feature a more outcome-focused or process-focused approach. For instance, the object of an outcome-focused evaluation could be the methods and tools used and the final service as such, whereas a process-focused evaluation would investigate the relations between the final service outcome and the process that created it, as well as assessing the collaborative performance during the process.

Therefore, this master´s thesis suggests that the approach taken by Foglieni, Villari & Maffei (2018) could be applied to co-design evaluation as well, dividing between service evaluation and evaluation of the co-design process in service development projects.

This master's thesis sees an opportunity for co-design evaluation during different stages of the process, with different evaluative aims. In fact, during the development and validation of the idea, the accent of the evaluation could be put on assessing the collaborative performance and the co-design process, while after the implementation the metrics could focus on evaluating the effects of co-design on the service performance.

This distinction in focus highlights a further differentiation in the time of the evaluation. In fact, metrics can be implemented before or after the implementation of the service. Once more, this highlights the division of the measurements into

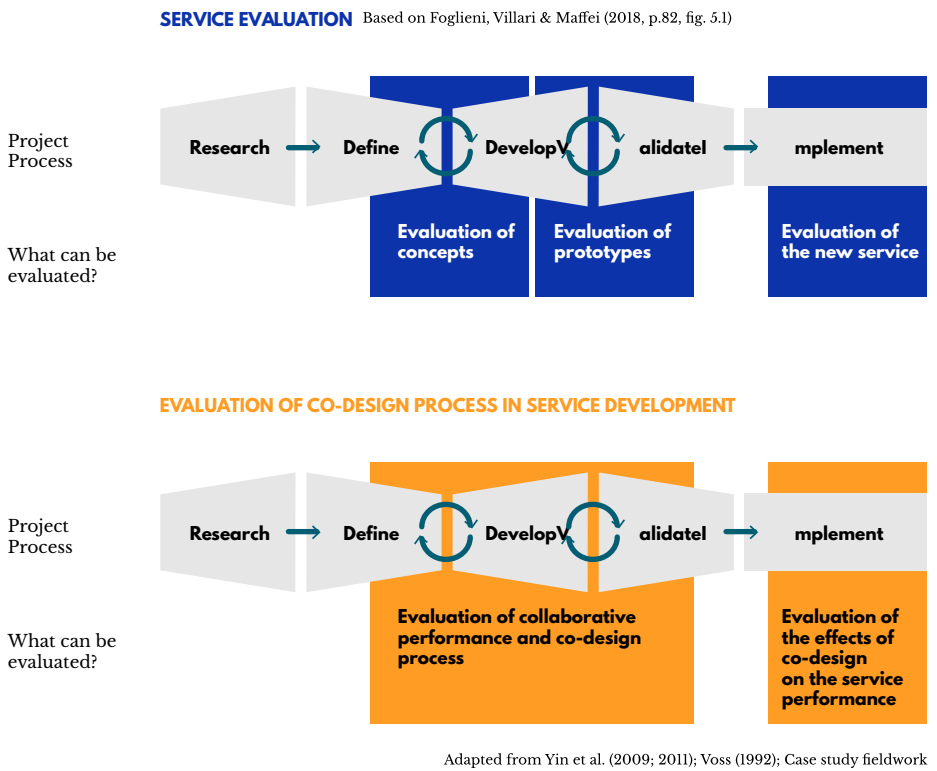


Figure 24. Service evaluation opportunities (Based on Foglieni, Villari & Maffei (2018, p.82, fig. 5.1) and evaluation of co-design process in service development opportunities.

Adapted from Yin et al. (2009; 2011); Voss (1992); Case study fieldwork

two categories: the measurements aimed at assessing outcomes after the implementation, and the measurements aimed at assessing the process before the implementation.

According to Voss (1992), the majority of the measurements deal with measuring the "performance of the innovation" (Voss, 1992, p. 44), meaning that they assess the outcomes and the results of the service. He also underlines how companies should be aware of the "determinants of success and failure in the process of service innovations" (Voss, 1992, p.44). This brings attention to the fact that most measurements are used in retrospective and thus serve to assess the results of a process that already happened. This is useful in terms of drawing conclusions on the service results and see how it is performing after implementation, as well as collecting best practices and learnings to apply in future projects. Furthermore, outcomes-focused measures cannot be used for informing changes to the process as it develops. In addition to producing changes *in itenere*, process-focused measures can also offer the opportunity to reflect on the service results looking at them through the lenses of the process that led there. In this sense, process-related measurements are useful to raise awareness and collect data to advocate for co-design.
Combining the division between service evaluation and co-design evaluation together with the differentiation between results-focused and process-focused metrics generates a matrix of the evaluation opportunities in a service design project. In *Figure 25*, the service evaluation part is grayed out to highlight that it is not part of the scope of the thesis.
For the scope of this thesis only the part concerning co-design evaluation is taken into account to further develop an evaluation framework. Future research opportunities might reside into into complementing the framework with metrics related to service evaluation.
Figure 25 represents an opportunity to synthetize and attribute specific evaluation criteria to the overarching categories represented by the "evaluation of collaborative performance and co-design process" and the "evaluation of the effects of co-design on the service performance". In order to do so, the project process needs to be expanded and more carefully articulated. From the interviews conducted in the

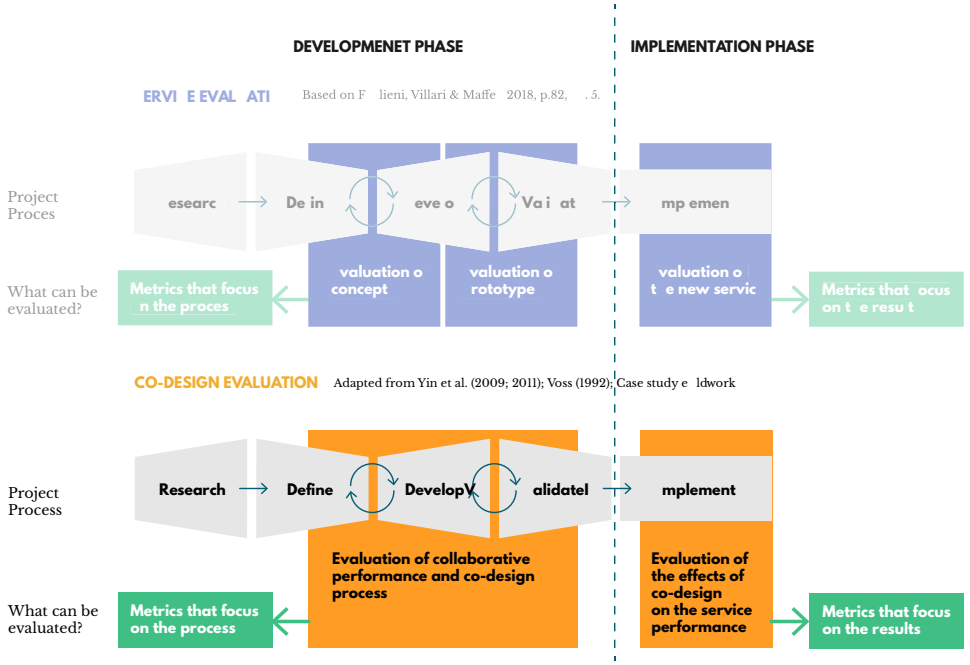


Figure 25. Evaluation of co-design process in service development opportunities: outcome-focused and process-focused metrics.

research phase, the service design process emerges as divided into two parts: the design phase and the execution phase, which corresponds with the division between before implementation and after implementation. Both phases are comprehensive of three "sub-steps" each. The design phase consists of "identify problems and needs", "generate ideas", and "validate problem-solution fit". The Execution phase consists of "service implementation", "iterate product-market fit", and "scaling the service". The line between the design phase and the execution phase is drawn by the executive buy-in. The design phase is renamed development phase and the execution phase is renamed implementation phase to translate the terms emerged during the Buddyschool research to hint for a scalable outcome.

The mapping of service design evaluation opportunities identified above on the service process represents the structural core of this evaluation framework. During the development phase, different metrics can be applied in different steps of the phase in order to evaluate the co-design process and the collaborative performance. Similarly, during the implementation phase, other metrics can be applied to the sub-steps of the phase in order to evaluate the effects of co-design on the service performance. These passages illustrated the thinking process and the reasoning behind the creation of the

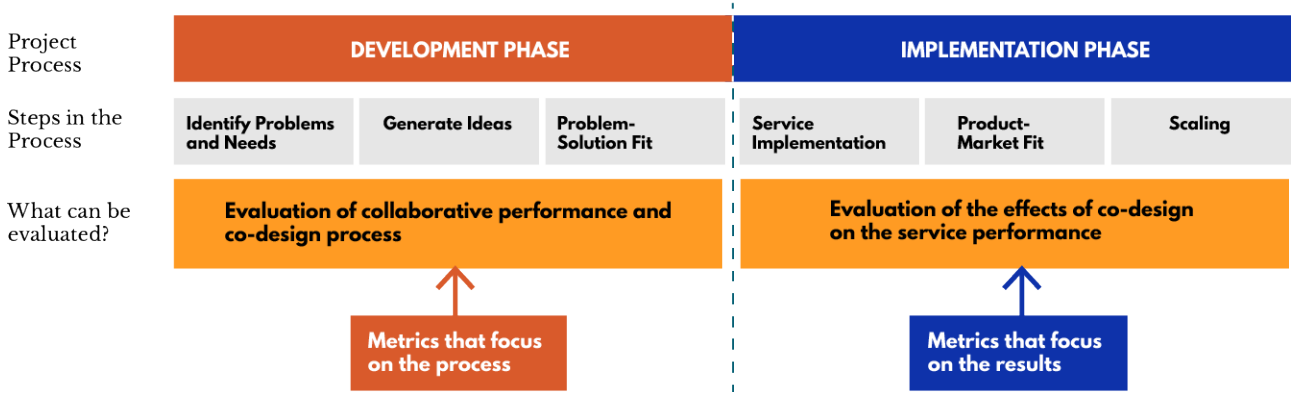


Figure 26. Core of the framework showcasing the project process divided in the two phases, the steps of the process and opportunities for evaluation.

the core structure of the framework. After the realization of the architecture, metrics gathered through the literature review, the experts' interviews and the case study fieldwork are placed inside the structure to actually create the framework.

The value that the framework adds to previous research is compiling a synthesis of evaluation criteria based on the literature and on field research. The measurements and metrics are divided in each step of the process, highlighting an opportunity to assess qualities specific of a certain phase of the process. Some metrics can be applicable to more than one process´ step, and several are part of both the development and implementation phases. Nevertheless, the metrics in the development phase are generally considered as leading indicators, whereas the metrics in the development phase are considered as lagging indicators. A leading indicator is an indicator of performance that might influence the future success of the service, hence the belonging to the development phase. Leading indicators are important to keep track of because they offer an overview on how the process is going in order to adjust the trajectory *in itinere* (Manuele, 2009). On the other hand, lagging indicators express past performances and measure the success after the implementation (Manuele, 2009). In this sense, the metrics belonging to the implementation phase are lagging indicators as they measure the effects that the process had on the service results.

The next section offers an overview of the metrics present in the framework and explains their grouping in more detail.

6.2 Design Proposal

In addition to following the structure presented in the previous section, the framework's content is categorized according to six different "thematic areas". These represent high-level elements and themes observed in the fieldwork and in the literature review. The areas are: participants and actors, ideas, concepts and prototypes, teamwork, learning and changing, values and objectives, efficiency and effectiveness. The six different thematic areas help in organizing the framework and providing the opportunity to cluster the metrics handling common topics together. The goal of organizing the metrics according to the clusters is enabling a more clear and simpler approach to the framework.
Each thematic category gathers different metrics together. Each metric is presented with a title, a description, and a link to the resources from where it is adapted from. The metrics' descriptions are mainly in the form of questions because the main goal at this stage is sparking further thoughts and reflections on the specific metric.
Future research should be directed towards a more clear and specific definition of the metrics.

Before providing an overview of the metrics present in the framework, three clarifications have to be made.

Firstly, the following evaluation framework is an attempt to crystallize the findings gathered from the literature review and the fieldwork of this thesis into a structured work. This means that the following visualization represents a general evaluation framework that needs to be applied to a specific project. The value that the general framework adds to the previous research is compiling a logical list of metrics with their resources, ready to be adapted and used in specific projects.

Secondly, because of the general nature of the framework, the terminology adopted remain on the general level as well. It will be up to the users to articulate the terminology in a more specific way, depending on the nature of the project, the actors, the purpose of the evaluations, and the results that the users are seeking. For example, the general framework chooses to use the general term "co-designers" to refer to the different actors in the co-design process. In reality, the term "co-designers", as in the actors that co-design, involves different roles and figures: co-designers as users taking part in the development process, co-designers as in the design team, co-designers as in the organizational actors, and so on. The goal behind choosing to implement a general term is aiming at maximizing the potential of the framework and leaving freedom to the users to articulate it in the ways that are most valuable for them in a given moment. Some of the metrics refer to more specific terminology (e.g. "users") because they are thought for a peculiar group, but mostly the metrics address the general group of "co-designers". When using the framework, it is the users who decides which particular group within the co-designers should they address. In this way the metrics will offer different opportunities for reflection depending on which actors are chosen. For instance, the metric "Changes in Mental Model – Are the co-designers implementing new ways of working learnt through the co-design process?" transforms its meaning according to the actors selected as focus. Consequently, the learnings developed through the framework are different: if one decides to focus on investigating the changes in mental model s coming from the organizational side, this metric will provide help, for instance, in understanding if there were changes in internal processes, ways of working or managerial and employees' mindsets. On the other hand, if one would like to articulate the metric thinking about the users taking part in the co-design process, the reflections will be directed towards new ways of thinking and doing of the specific users that participated in the development. This would give the freedom to the users of the framework to explore different implications and effects according to different actors that took part in the process.

Limitations and suggestions for future research on this topic is presented in the next chapter, *7. Discussion*.

Lastly, the framework only focuses on co-design evaluation metrics, as explained before. Even though the actions of evaluating the service and evaluating the co-desing process that led to the service could potentially have several metrics in common, the way the metrics are articulated is different. Specifically, co-design evaluation metrics focus on investigating the relation of the elements with the process, whereas service evaluation assessments focus on the elements themselves. For instance, the thematic category of "ideas, concepts and prototypes" can be present both in a co-design evaluation framework and in a service evaluation framework. Even if the thematic area might be the same, the metrics belonging to it would be different. Specifically, the metrics in service evaluation would focus on the elements and their qualities, therefore on the quality of the ideas, the number of prototypes, the selection criteria of the concepts, etc. On the other hand, in a co-design evaluation framework the metrics included in the thematic area of "ideas, concepts and prototypes" would put the elements in relation with the process, for example investigating if the co-design process (and therefore the fact that people with different backgrounds and roles were sitting at the same table) helped in generating different ideas and if these ideas where more accurate, more straightforward, more polished, sharper, better, more intrigant, more thought out because of the collaborative process that jointly brought them about.

In this sense, if it is true that the last section provided an overview of the core structure of the framework dividing in development phase with process-focused metrics and implantation phase with outcome-focused metrics, it is also true that co-design evaluation is infused with an overarching interest towards the process and its effects and impacts on the service results. Therefore, all the metrics presented in this framework investigate to some extent the relation of an element with the process.

Figure 27 offers a preview of the framework for co-design evaluation in service development projects. A bigger representation is featured in the next page.

The rest of this section provides an explanation of each metric.

Figure 27. Thumbnail of the whole framework. The next page features a whole page representation.

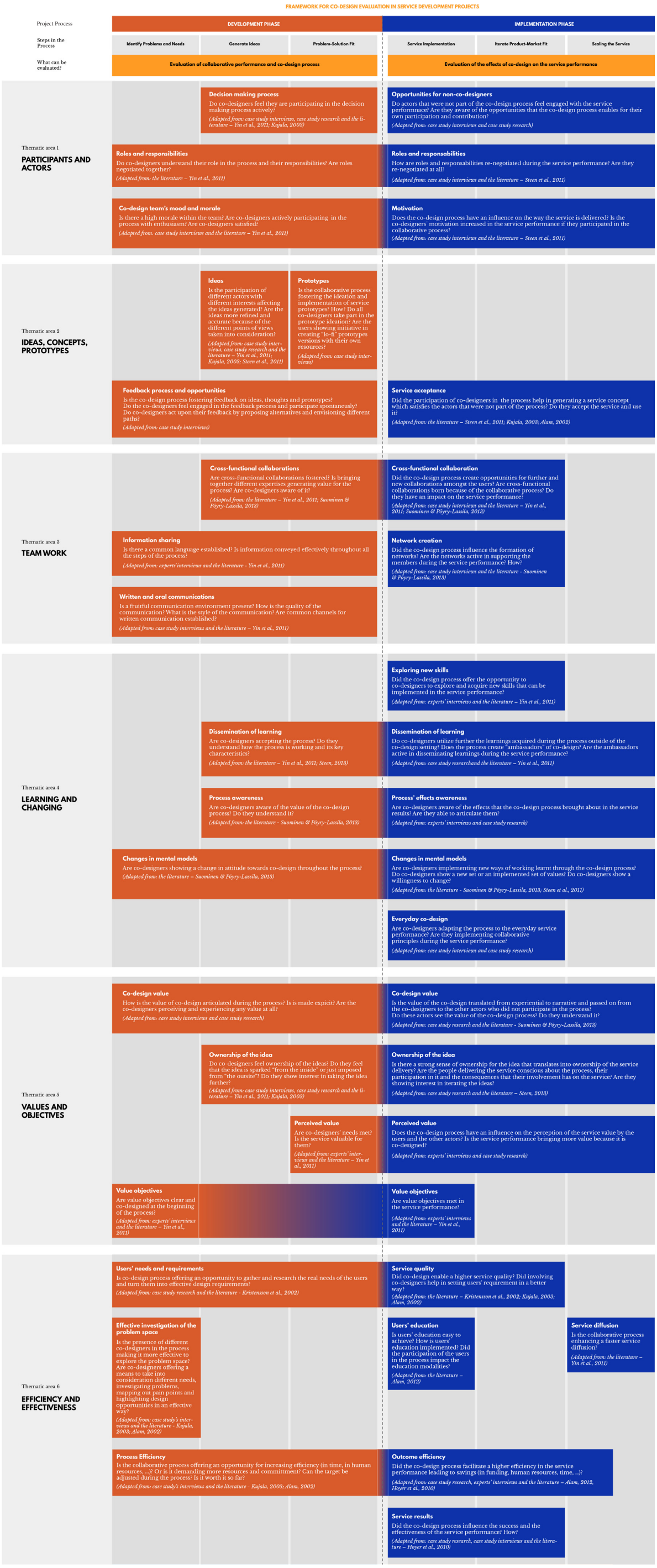


Figure 28. Whole page representation of the co-design evaluation framework proposed in this thesis.

1. Participants and Actors

The first thematic area is "Participant and Actors" gathering the metrics referred to the experience of the people taking part in the process.

- Decision Making Process.

The aim of this metric is investigating the participation of the co-designers in the decision-making process and assessing their opportunities, but also their interest in being actively involved.

- Opportunities for non-co-designers.

The metric, belonging to the implementation phase, investigates the knowledge and awareness of non-co-designers on their own possibilities and opportunities for contributing the service further during the service performance.

- Roles and Responsibilities.

This metric presents the opportunity to reflect on the roles of the different actors, their implications, and their responsibilities. The part of the metric articulated in the development phase aims at assessing the situation as it develops, offering a chance to have a more thorough discussion about roles during the process. On the other hand, in the implementation phase, the metric highlights how and if the roles are re-negotiated during the service performance.

- Co-design team's mood and morale and Motivation.

The first part of the metric assesses the mood of the co-design team during the co-development process with the goal of opening discussion in case the morale would not be positive. The second part of the metric explores the correlation between participating in the process and presenting a positive motivation during the service performance.

2. Ideas, Concepts and Prototypes

The second thematic area collects metrics aimed at assessing ideas, concepts and prototypes, specifically investigating the relationships between implementing a co-design approach and the quality of the ideas, as well as the feedback process.

- Ideas.

This metric investigates if the collaborative nature of the idea generation represented a positive aspect.

- Prototypes.

Similarly to the Ideas metric, it explores the possibilities of co-design fostering the ideation and implementation of prototypes.

- Feedback Process and Opportunities.

This metric addresses the possibility of the collaborative process to bring about a more open-minded response to the feedbacks by the participants, supporting it.

- Service Acceptance.

This metric offers an opportunity to reflect on the level of satisfaction of the users and if it was connected with the choice of implementing a co-design approach.

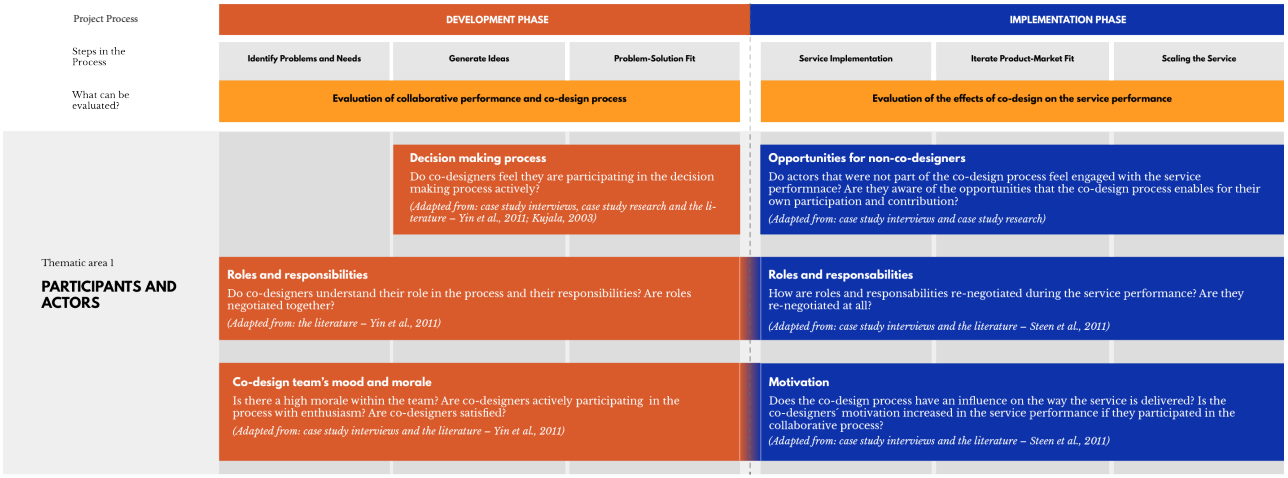


Figure 29. Thematic area 1: Participants and Actors.

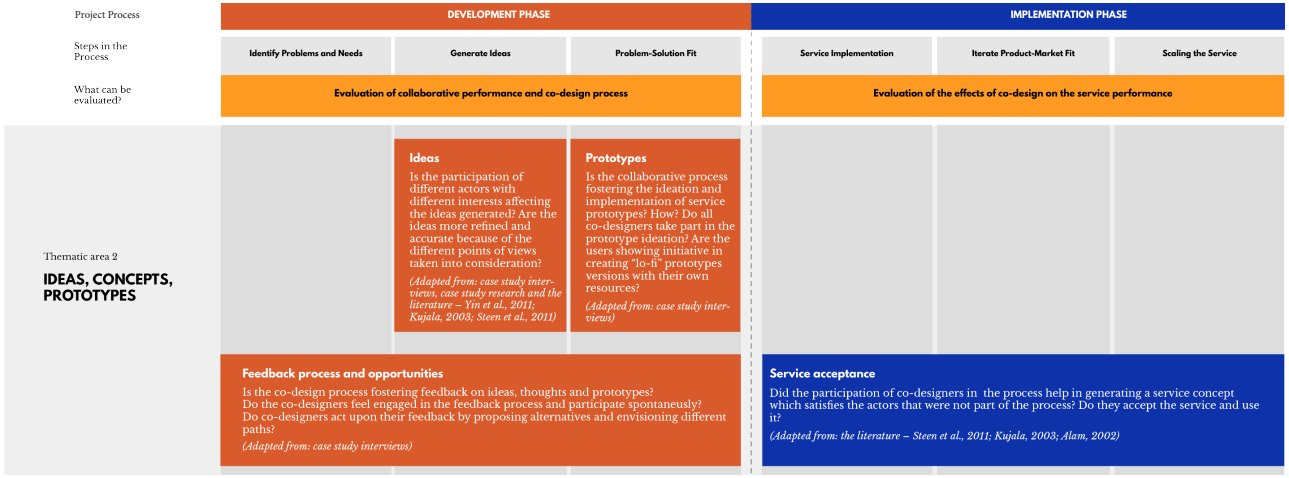


Figure 30. Thematic area 2: Ideas, Concepts, Prototypes.

3. Teamwork

The third thematic area "Teamwork" focuses on the collaborative aspect of working together within and across teams.

- Cross-functional collaborations.

The first part of the metric in the development phase aims at reflecting in itinere about the creation of such collaborations and their value, as well as keeping track of them. The second part of the metric in the implementation phase focuses on understanding the implications that the new cross-functional collaborations had on the service performance and their impacts on the execution phase.

- Information sharing.

The metric in the development phase aims at fostering reflections on the creation of a common language for communication while the process develops.

- Written and oral communication.

If the previous metric focused on assessing if a common communication style is achieved, this metric focuses on understanding if communications are effective. As part of the development phase metrics, it provides the opportunity to reflect during the process and make adjustments *in itinere*, if needed.

- Network creation.

This metric in the implementation phase focuses on understanding the correlation between the formation of networks amongst and beyond the co-designers and the co-design process.

4. Learning and Changing

The fourth thematic area collects together the metrics aimed at assessing changes in ways of doing and thinking, as well as evaluating the learning emerged from the co-design process.

- Exploring New Skills.

The metric investigates if the co-design process offered opportunities for co-designers to learn and discover new skills and capabilities. As part of the implementation phase, the metric offers an overview of the potential personal accomplishments and growth of the co-designers.

- Dissemination of Learning.

The first part of the metric in the development phase explores the awareness of co-designers on the principles of co-design and the ways the process is working. The second part of the metric in the implementation phase focuses more on assessing if the process-related learnings are used during the service delivery, and if the process created "ambassadors" who would spread the co-design awareness to other actors who did not participate in the co-development.

- Process Awareness and Process' Effects Awareness.

The first part of the metric in the development phase offers an opportunity for reflection on the level of awareness of the actors participating in the co-development in order to act upon it. The second part of the metric in the implementation phase focuses more on understanding the

awareness of the actors on the effects brought about by the collaborative process on the service results.

- Changes in mental model.

In the development part, the metric keeps track of the changes in mental model by the co-designers, whereas in the implementation phase it assesses the changes observed during the service performance, symbolized by the implementation of new ways of working.

- Everyday Co-design.

The metric focuses specifically on understanding if the co-design principles are translated and adapted to the everyday co-design setting, fostering a co-design culture across boundaries.

5. Values and Objectives

The fifth thematic area gathers the metrics that focus on values and objectives of the co-design process.

- Co-design Value.

The first part of the metric in the development phase addresses the value of the process during the co-development phase, especially the opportunity for making it explicit to the co-designers. On the other hand, the second moment of the metric in the implementation phase focuses on assessing if a translation from experiential to narrative knowledge happened, in other words if the knowledge and the awareness has passed on from the co-designers to the actors not

Figure 31. Thematic area 3: Teamwork.

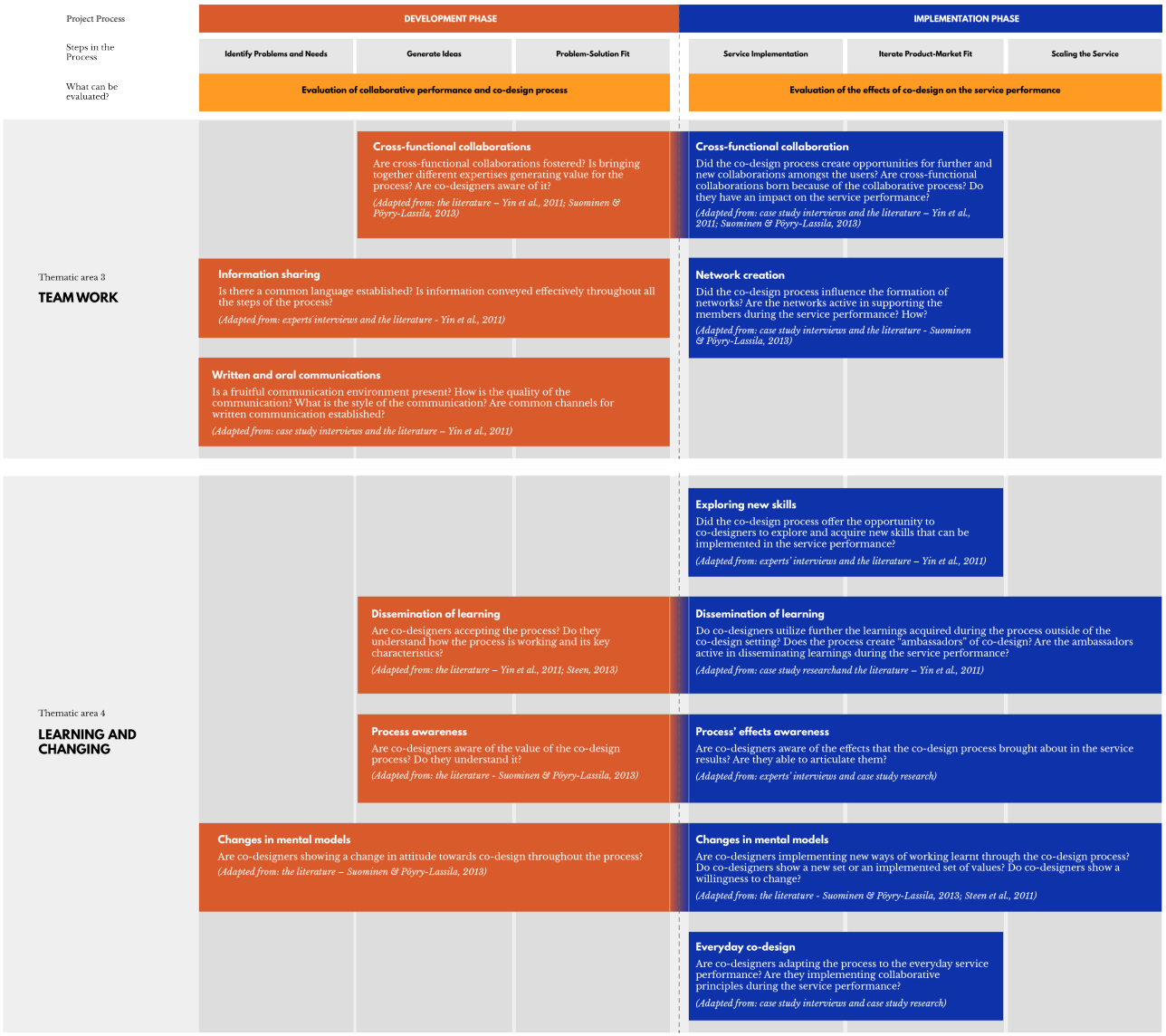


Figure 32. Thematic area 4: Learning and Changing.

participating in the process.

- Ownership of the Idea.

The first phase of the metric investigates the value through asking questions about the relationship that the co-designers feel with the idea. The second part of the metric investigates if the sense of ownership is translated across the service implementation and which implication this has on the service performance.

- Perceived Value.

The first part of the metric focuses on understanding the value of the service for the users. Doing it during the design phase gives the opportunity to adjust the trajectory if needed. The second part of the metric explores the relation between the perceived value and the fact that the process was collaborative.

- Value Objectives.

Firstly, the metric aims at reflecting on the collaborative creation and setting of the value objective for the service. In the second phase, the focus is reflecting on the actual implementation of the value objectives.

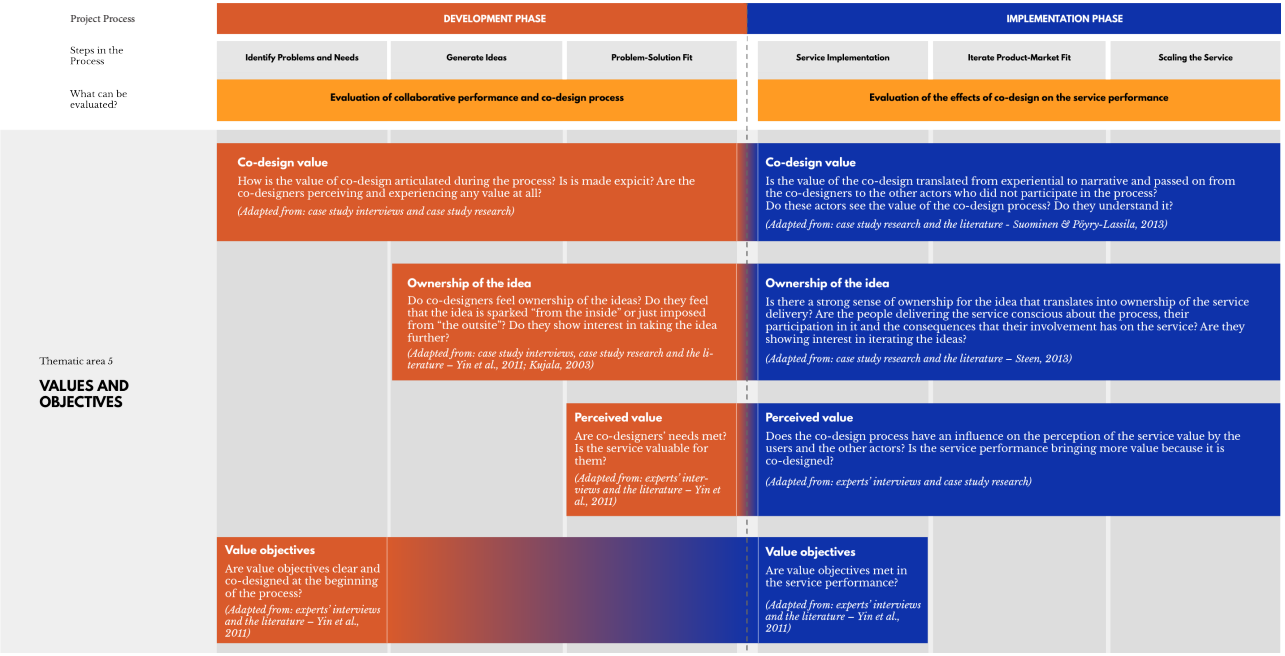


Figure 33. Thematic area 5: Values and Objectives.

6. Efficiency and Effectiveness

The sixth and last thematic area offers an overview of metrics that deal with efficiency and effectiveness of the process.

- Users' Needs and Requirements and Service Quality.

During the development phase, the relation between the co-design process and the creation of effective users' requirements is explored. In the implementation phase, the focus is shifted on the relation between the service quality and the co-development process.

- Effective Investigation of the Problem Space.

The metric offers the opportunity to reflect on the value that having different backgrounds within the co-designers might bring to the research phase where the problem space is investigated.

- Users' Education.

The metric explores the relationship between the co-design process and the modalities in which users' education is achieved.

- Service Diffusion.

The metric investigates the link between service diffusion amongst users and the fact that the service is the result of a co-development process.

- Process Efficiency and Outcome Efficiency.

The first part of the metric in the development phase gives the opportunity to reflect on the usage of resources during the process, whereas the second part of the metric in the implementation phase puts the accent on how the resources

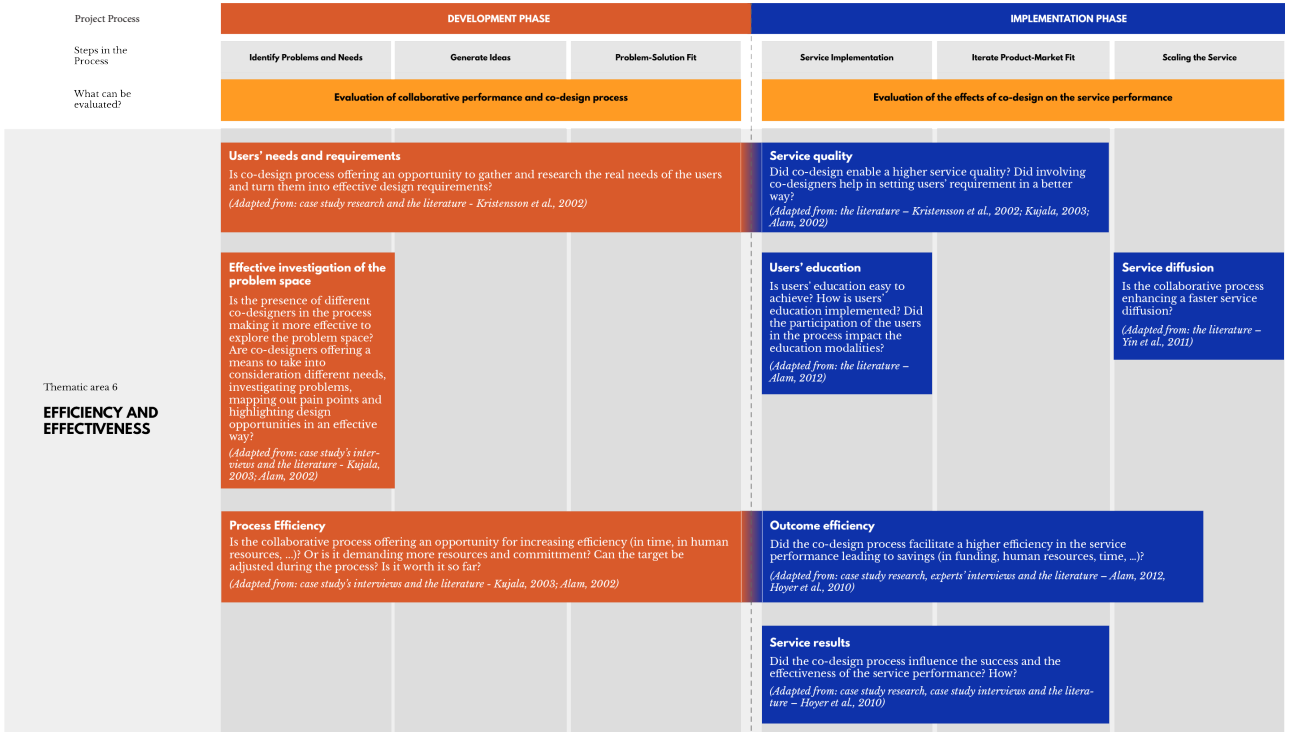


Figure 34. Thematic area 6: Efficiency and Effectiveness.

are being used during the service performance, if leading to savings or extra costs.

- Service Results.

The metric explores the possible relation between the choice of a co-development approach and the success and effectiveness of the service performance.

Rather than evaluation as an outcome-focused activity, the framework wants to promote evaluation as a means for creating the conditions for co-design's awareness to grow, learning from the process in order to understand what needs to be changed and what works well. In this sense, the framework presented in this master's thesis must be read with a reflective purpose above all. Therefore, the descriptions of the metrics are left in a question format, to offer food for thoughts on a specific project. The aim of the framework is not sparking a negative or affirmative evaluation judgment, e.g. "this specific project is better because it has been co-designed" as the current framework does not offer the tools and opportunities to conduct such evaluation, since it lacks in offering meters of comparisons with other projects. This issue will be touched upon in the next chapter, *7.1 Suggestions for Future Research* and *7.2 Limitations of the Research*.

In order to validate the framework and its application to specific projects, the Buddyschool case study was taken as pilot. In the next section an overview of the framework applied to Buddyschool is presented, as well as considerations on the metrics applied to specific projects.

6.2.i Applying the Framework to Buddyschool Case Study

As stated before, while the framework is general and gathers different metrics, it has to be applied to specific projects. This means that a selection of the relevant metrics must occur case by case, as not every item could be applicable within any project. In this research the general framework was applied to the Buddyschool case study. This underlines how, overall, the literature review played a key role in this thesis, even though it was not initially expected. In fact, this research elaborates a framework for evaluating co-design and then applies it to the case study

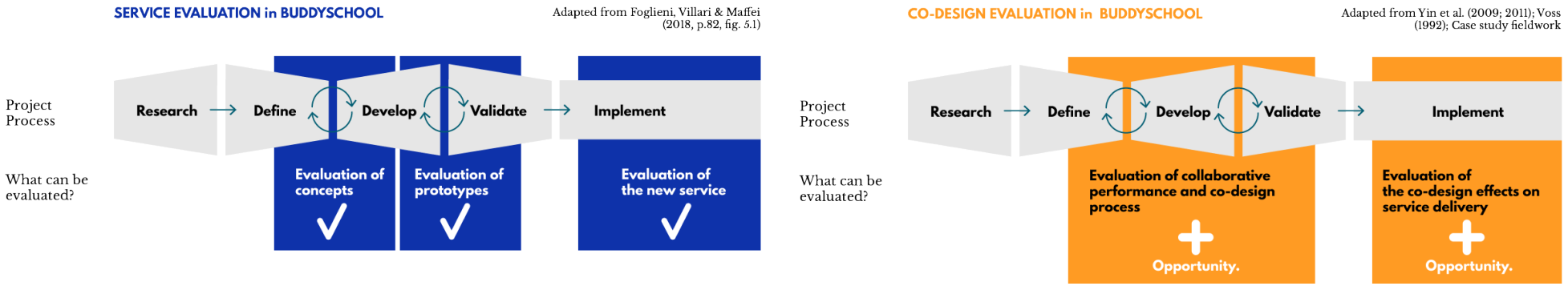


Figure 35. Service evaluation and Co-design evaluation opportunities in Buddyschool

of Buddyschool, meaning that the proposal is potentially applicable to other services as well with the right considerations. The topic is also touched upon in the "Future Research" section in the next chapter, *7. Dicussion*.

In order to apply the framework to the case study of Buddyschool, the first step was taking into consideration the evaluation process and practice in the case study. Even though this master's thesis focuses on co-design evaluation, as highlighted above, a brief overview on the service evaluation in Buddyschool must be presented in order to gain contextual knowledge of what typology of evaluation was carried out during the process.

Three moments can be observed when service evaluation was implemented in the Buddyschool design process. Firstly, evaluation was used to assess the different ideas and service concepts generated through co-design activities. Evaluation was used to prioritize which ones were the most valuable and should move to next phase of experimentation. In this phase, feedback was gathered from the co-design team as well as from a group of "young designers", who commented the ideas and were also asked to report back possible feedback received from their peers. The second evaluation moment is represented by the evaluation of service prototypes. During the experiments, different versions of Buddyschool were tested in real life environments in order to assess the feasibility and how the different options with different characteristics were working. Lastly, the third evaluation moment was after the service implementation. This feedback session aimed at gathering data and opinions on the delivery and the effectiveness of the service by the stakeholders involved. The primary data was anecdotes from teachers and numerical information on students' performances and the number of schools subscribing to the service. Even though they are presented as "evaluation moments", the process was not linear and sometimes different evaluations overlapped. In addition, the evaluations were not crystallized in one step of the development process, but they were continuously implemented throughout the whole phase in which they were deemed important. Therefore, the framework of service evaluation for new services (evaluation of concepts, of prototypes, of the new service) as explained by Foglieni, Villari & Maffei (2018, p. 82) can be observed in the actions adopted during the Buddyschool design process. *Figure 35* showcases service evaluation and co-design evaluation opportunities in Buddyschool.

Understanding the service evaluation carried out in Buddyschool is useful to understand how co-design evaluation could be articulated, and to craft a clear journey of the actions occurred in each phase of the process. As presented before, a possibility of evaluating co-design lays in the evaluation of the collaborative performance during the development and validation of the concept. The assessment of this performance *in itinere*, as undelined by Voss (1992), would provide means for reflecting on the process and improving it as it develops. The only parameter used for evaluating such performance by Migrant Youth Helsinki was the number of people participating during the workshops and if their participation was continuous. Since there was no formal evaluation process in place, proposing now to evaluate the collaborative performance must be seen as a hypothesis that remains on the retrospective level. The second opportunity for evaluating service design can be found in evaluating the impacts of the co-design process in the service performance. In other words, assessing which are the effects that the co-design process has on the implementation of the service and, if possible, determining to which extend they contributed to the service performance.

After understanding the service evaluation process and the co-design evaluation opportunities for Buddyschool, the design journey must be divided into actions according to the six overarching steps of the service process. Understanding which actions in the process represent each step is essential in mapping the journey and recognizing which metrics could be useful for each phase in this specific case. This means applying the steps in the process mapped in the framework architecture according to the specific process that occurred in the Buddyschool case study.

Before focusing on the specific metrics and indicators, a selection of the main actors to take into consideration in the framework must be carried out. This example decided to highlight the role and the opportunities of the teachers that participated in the co-development process, namely the Teachers Co-Designers.

Project Process	DESIGN PHASE			EXECUTION PHASE		
Steps in the Process	**Identify Problems and Needs**	**Generate Ideas**	**Validate Problem-Solution Fit**	**Service Implementation**	**Iterate Product-Market Fit**	**Scaling the Service**
What happened in Buddyschool development?	Workshops exploring thematics: education, working life, social landscape.	Workshops where adult co-designers generated ideas on the theme of edcuation.	1. Feedback from Young co-designers on the ideas. 2. Young and adult co-designers tested the ideas with their networks. 3. Service experiments in the schools.	Schools start implementing and using Buddyschool. Product owner provides help in setitng up the service and then it is up to the teachers.	1. Teachers can deliver the service with the modalities they see fit for their school, although the core structure remains the same. 2. Product owner does interviews and questionnaires with the teachers to gather feedback on results.	1. From one school to many schools (quantity). 2. From a small scale to a big scale in the same school (size).

Figure 36. Co-development process in Buddyschool mapped out against the core structure of the Framework.

It must be underlined that the aim of applying the framework to Buddyschool is not running an actual evaluation, but only translating the metrics to the specific service and case study in order to validate the framework's architecture and ways of working. No evaluative judgment is featured in the framework, but the metrics are only adapted in order to validate the application of a general framework to a specific project.
In addition, the framework application to Buddyschool case study is the result of personal reflections, as no validation from the Buddyschool team's side is present. This means that the choice on which metrics and actors to highlight was entirely taken by the researcher.
This issue is discussed in the next chapter, *7. Discussion*, in the section "Limitation of the Research".
In the following sections, a brief overview of the framework's metrics applied to the Buddyschool case study is presented. A selection of relevant metrics according to the participants selected is made beforehand. Additionally, the metrics pertinent to the development phase are explained by taking into consideration the added value they could have brought to the process as a tool for making changes *in itinere* in retrospective.

Hereafter, the thematic areas applied to Buddyschool are explained and an overview on the metrics is presented.

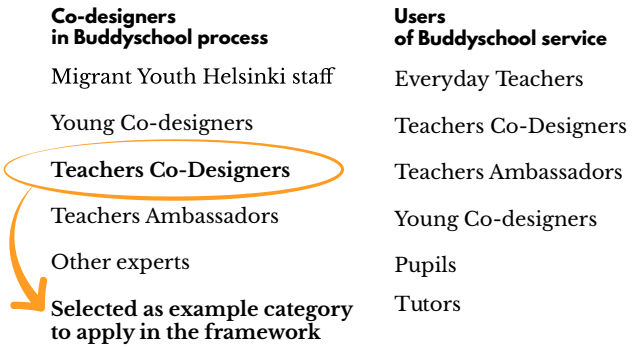


Figure 37. Actors in the Buddyschool process with highlighted the selection of who to take into consideration in the framework application.

1. Participants and Actors

In the first thematic area, "Participants and Actors", metrics that deal with the feelings, the motivations and the emotions of the participants in the process are presented. Specifically, Teachers Co-design are taken into consideration. Addressing their involvement in the process by investigating their feeling of inclusion in the decision-making process and their mood and morale throughout the process is seen as a key point for potentially make changes during the development phase. In addition, the translation and transfer of the teachers' motivation during the process to the service performance is taken into consideration.

Moreover, one metric focuses on roles and responsibilities during the process. Understanding the take of Teachers Co-Designers on their roles and their involvement in the process is useful for investigating the ownership of the service idea and the service delivery after implementation.

2. Ideas, Concepts, Prototypes

Since a service evaluation has already been implemented during the Buddyschool process and after the service implementation, the thematic area "Ideas, Concepts, Prototypes" in the co-design evaluation focuses more on investigating the relation between the co-design process and the different ideas and prototypes. Particularly in this case, analyzing the involvement of Teachers Co-Designers in the process and what effects that had on the idea generation and the service prototypes phase is the focus of the current metrics selected. In addition, the engagement of the Teachers Co-Designers in the feedback process is addressed in order to potentially make changes to the process during the development.

3. Teamwork

The thematic area of "Teamwork" is adapted to the Buddyschool case study by taking into consideration metrics aimed at reflecting on the creation of cross-functional collaborations and the style of the communication with other

participants during the process. Investigating the typology and the means of communication is useful for understanding if the values are uncovered during the process with a language and a way of communicating that is understandable and relates to the Teachers Co-Designers. Furthermore, the opportunities for cross-functional collaborations are researched during the design phase, as well as investigated after the service implementation, together with their ties to the co-design process. The considerations for the application of these metrics come mainly from the interviews with the Migrant Youth Helsinki Staff.

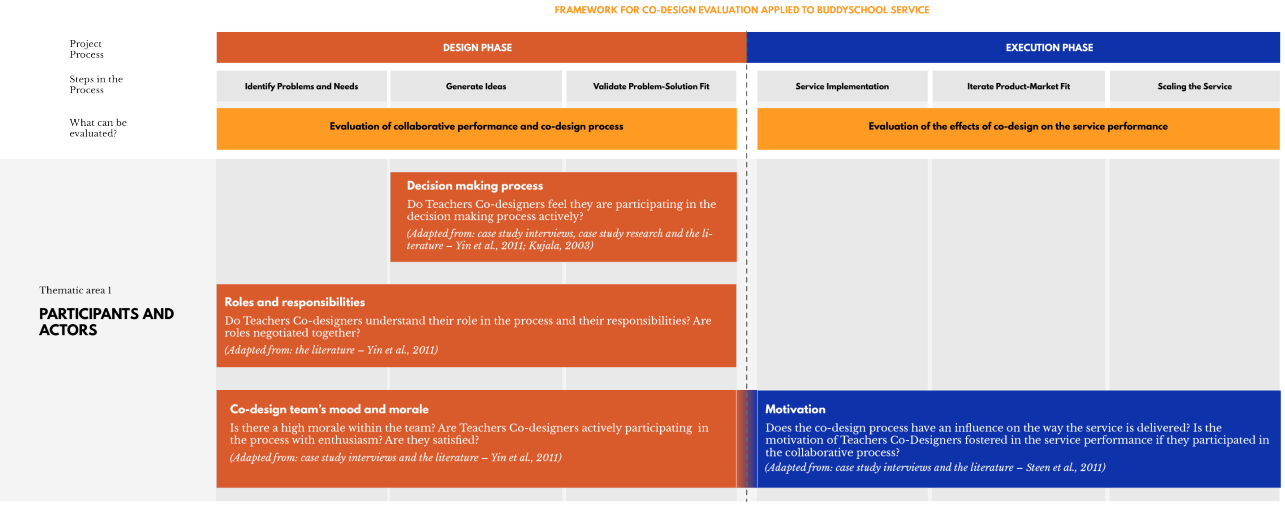


Figure 38. Category 1: Participants and Actors, focusing on Teachers Co-Designers in Buddyschool.

4. Learning and Changing

One of the most interesting areas to investigate in Buddyschool is "Learning and Changing". This applies to Teachers Co-design as investigating their attitude during the process and after the implementation of the service. Highlighting changes in mental models might be an interesting path to investigate a higher-level impact within the service. In other words, the goal is researching a different kind of impact that is not outcome-focused or effective-focused, but touches upon the involvement of the Teachers Co-Designers and their development as people and professionals during the process. This could also be investigated by observing if the learnings gathered during the process are now

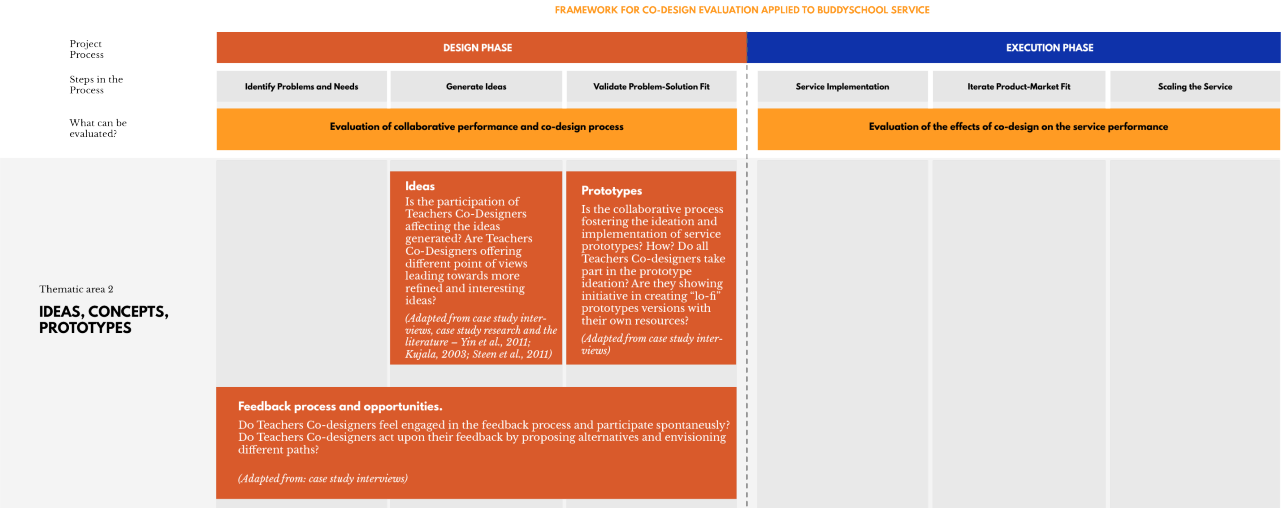


Figure 39. Category 2: Ideas, Concepts, Prototypes, focusing on Teachers Co-Designers in Buddyschool.

implemented and adapted to the everyday life of the Teachers Co-Designers: were they able to gather all the useful information and find ways to translate it into actionable points during the service performance?
This reflection is sparked mostly by the observations during the field visit in the school: if teachers are operating in a similar setting and environment than a co-design activity session, could the benefits of participating in a co-design process go beyond the benefits brought about by the development of a more valuable service? In other words, could the Teachers Co-Designers have used the opportunity to collect further learnings on the dynamics and the process in order to (implicitly or explicitly) adapt them to their everyday job?

5. Values and Objectives

Building on the importance of the "Leaning and Changing" section, the "Values and Objectives" area partially represents the preliminary investigation aimed at assessing the changes. In fact, analyzing the relation between the Teachers Co-Designers and the process, investigating if the values embedded in it are explicit for the teachers, might be useful for then being able to research the translation of such values and ways of working in the service performance. In general, understanding and being aware of the process dynamics and the co-design values is

Figure 40. Category 3: Teamwork, focusing on Teachers Co-Designers in Buddyschool.

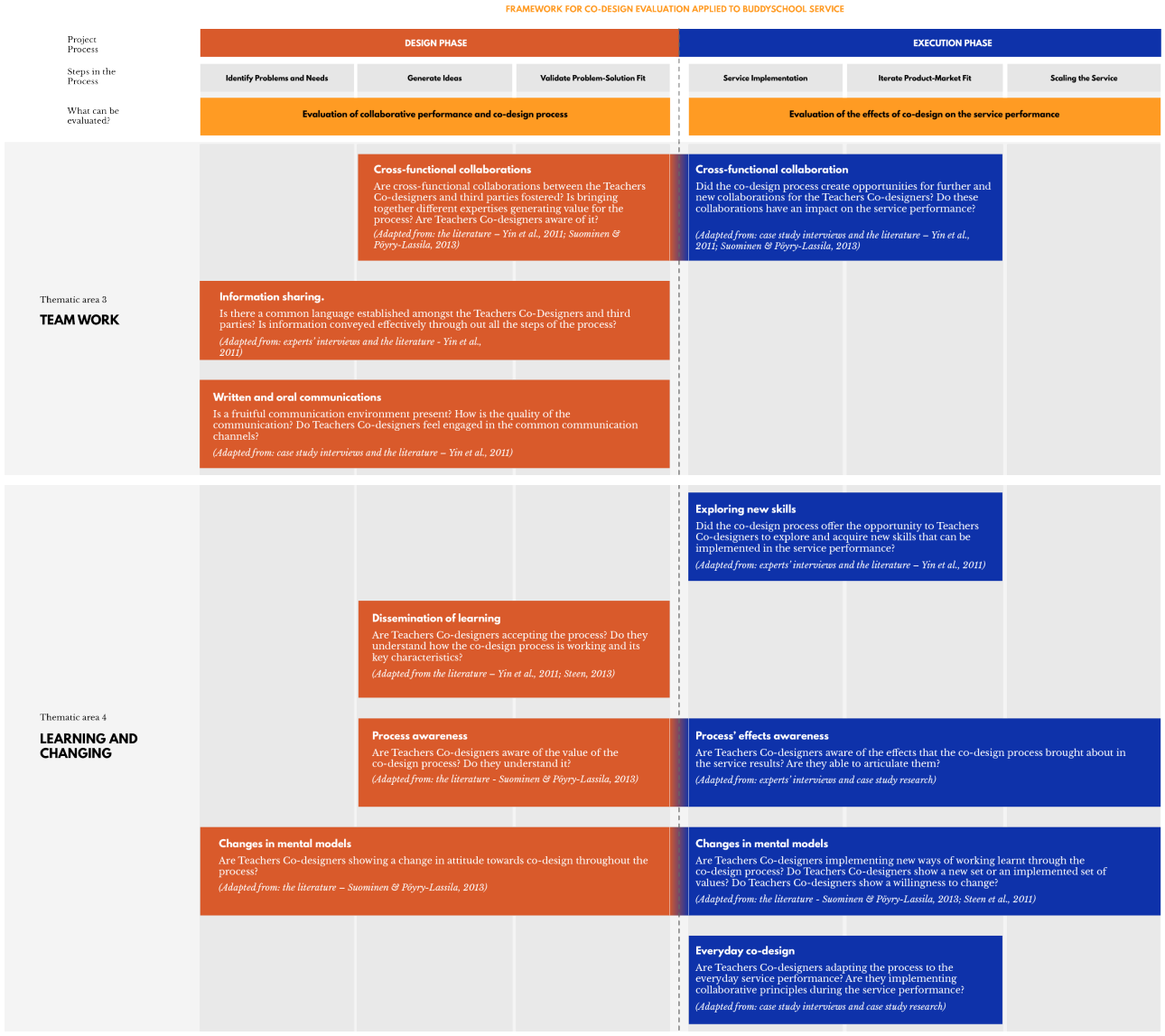


Figure 41. Category 4: Learning and Changing, focusing on Teachers Co-Designers in Buddyschool.

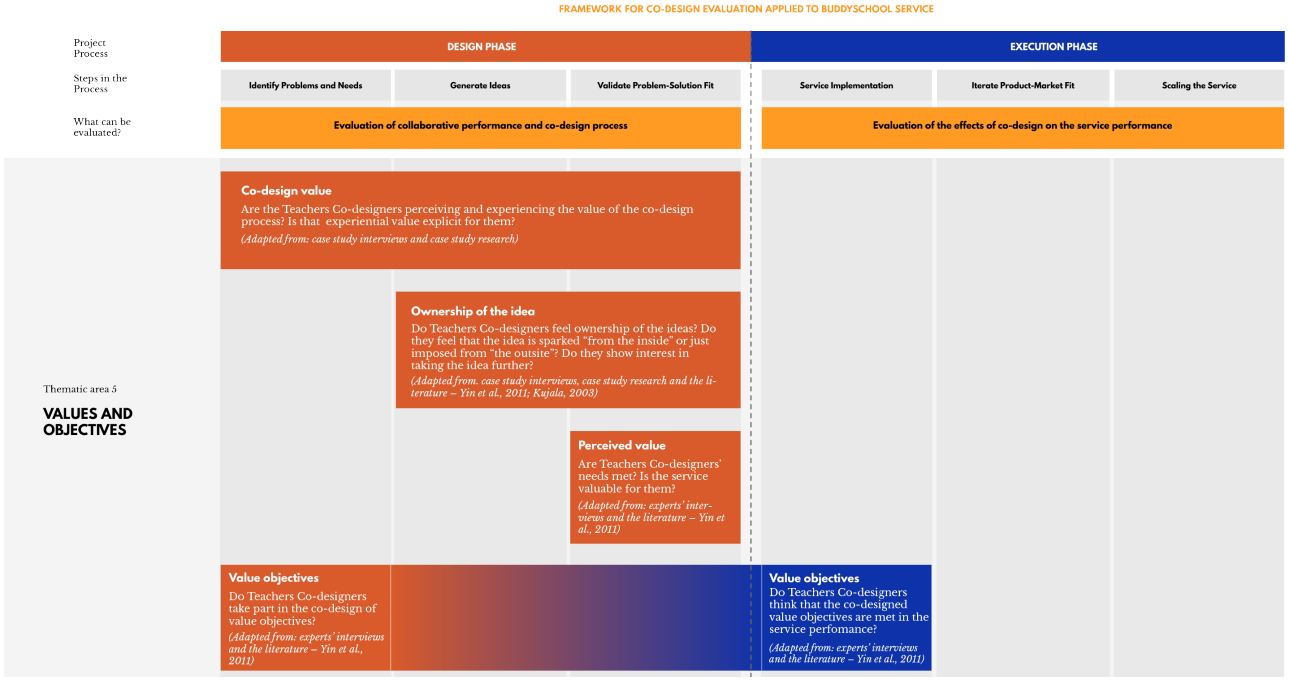


Figure 42. Category 5, Values and Objectives, focusing on Teachers Co-Designers in Buddyschool.

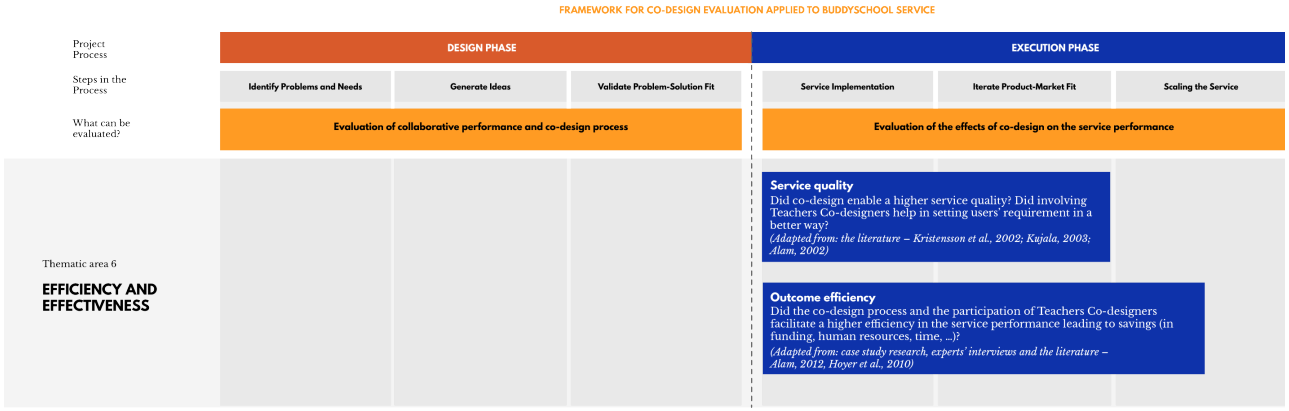


Figure 43. Category 6, Efficiency and Effectiveness, focusing on Teachers Co-Designers in Buddyschool.

regarded as a key factor for then being able to understand one's responsibility during the service performance, continuing the iteration and shaping the service according to the co-designed values objectives. Similarly to the "Learning and Changing" reflections, most of these consideration come from the school visit research.

6. Efficiency and Effectiveness.

Lastly, the "Efficiency and Effectiveness" category aims at highlighting the relation between the co-design process and a higher service quality, as well as efficiency. For the Teachers Co-Designers, realizing that the process is closely tied with a better, more efficient, and more effective service performance might be useful in order to better contextualize the importance of such process and maybe turn into Ambassadors advocating for the co-design process within Buddyschool. On the other hand, if metrics would not highlight a tight causal relation between the process and the service performance, this would be helpful to spark further thoughts on how to improve the process, fostering feedback for Migrant Youth Helsinki staff. In general, this area does not contain plenty of metrics and indicators because the aim was prioritizing reflection of the Teachers Co-Designers' experience during the process, the learnings that they achieved throughout the journey and how they

transformed and adapted them in the service performance. This underlines the flexibility of the framework as a tool: implementing an evaluation targeted to a specific project with the aid of this framework allows the users to decide which actors and which areas must be prioritized. In this way the learnings and the assessments are valuable and customized according to different needs and interests.

6.2.ii A Co-Design Evaluation Framework for Buddyschool: Conclusions

In general, evaluating the co-design performance and the process effects of the service performance is valuable for Buddyschool case study, and specifically for Teachers Co-Designers, because it allows them to understand their possibilities, their roles and their responsibilities for further iterating and developing the service. It also offers a chance for reflecting on the journey and the intrinsic value of the co-design process, learning and experimenting how to adapt and transform it in order to apply it as an everyday co-design process.

Although at the moment the framework does not present an opportunity for finding clear answers to the questions proposed in the metrics, the application to the case study of Buddyschool could still be useful and valuable for sparking further thoughts and reflections, drawing conclusions on the four-years project while getting ready for what will come next. In addition, it must be underlined how the evaluation framework would offer Migrant Youth Helsinki a flexibile opportunity for assessing different elements in the process and in the service performance. Since the staff from Migrant Youth Helsinki holds service design and design thinking capabilities, they would have the tools to make decisions on which metrics to investigate, even eventually shaping new and interesting metrics.

Furthermore, this evaluation framework applied to the Buddyschool case study would be useful in order to conduct further experiments in the schools, since it represents an organized way to potentially compare results gathered in different institutes in Helsinki.

In general, applying the general framework to the specific case study of Buddyschool must be seen as a validation exercise for the research. The current framework is only a first step in the direction of building a coherent, comprehensive and organized co-design evaluation framework. In the next chapter, the limitation of the research and suggestions for future research are discussed more in detail.

Chapter 07

Discussion

7.1 Suggestions for Future Research

This section will briefly present some possible further research areas concerning co-design evaluation and the design proposal featured in this master´s thesis, the co-design evaluation framework.

A first opportunity would be represented by the inclusion of service evaluation metrics in the framework, therefore transforming it into a hybrid work aimed at evaluating both the service and the co-design process behind it. The value of this hybrid framework would be to offer an opportunity for a holistic evaluation touching upon different areas. In fact, service evaluation is still to be considered an important part of the evaluation process and it provides the project with interesting data on the utilization of the service, as well as helping in making grounded choices during the whole process.

Indeed, before considering the implementation of a hybrid framework, the current one needs to be further developed in its components and in its structure. For instance, further research could be aimed at better defining the metrics presented in the current proposal. In addition, the framework could be expanded to encompass extra indicators in order to build more accurate and diverse opportunities for evaluation.

After more carefully establishing the metrics and the criteria for evaluation, future research might also focus specifically on indicators and measurements. At the moment the framework provides a list of metrics with a description that offers hints and cues on its utilization, but the current metrics do not feature any kind of qualitative or quantitative measurements to utilize in order to give a value judgement. Therefore, even though at the moment the main purpose of the framework is sparking thoughts and fostering reflections on a specific project, future research could delve into defining more accurate indicators and measurements.

In order to utilize the framework as a tool aiding in forming a value judgment, and not only as a means for reflection, future research might focus on creating a more structured architecture in order to allow comparisons between different projects. In fact, the research highlighted how evaluative practices need a meter of comparison in order to express a value judgement. While the current framework's goal is fostering reflections on a specific project by asking questions that might relate to that project, future research could investigate the possibility of enabling comparisons between similar projects with a different design maturity level. In fact, at the moment the data gathered through the framework is not offering an evaluative opportunity in terms of providing judgements (for instance, the metrics aimed at capturing the effects of the co-design process on the service performance cannot be currently used to prove that the effects are actually caused by the co-design process because of the lack of comparison with a situation where co-design was not developed).

Further research is also needed in the clarification and definition of the users referred in the framework. At the moment, the general category of "co-designers" is taken into consideration and the choice of which specific category of actors to address is left to the users of the framework. This could be addressed in future research, for example by already creating different proposals aimed at different actors' groups, or at least by articulating better these different opportunities.

Nevertheless, the current work adds value to the research work around the topic of co-design evaluation as it offers a way to organize different metrics based on the service outcome and the service development process. In addition, it helps in clarifying the benefits, the positive effects and the expected outcomes of such co-design processes.

Concerning the framework applied to Buddyschool´s case study, an opportunity for further development would be utilizing the framework as a tool to explore the service results and the implications that the co-design process had (and is having) on the service performance. Since the

project is soon coming to an end and it is ready to be fully passed on to the school system, the framework present in this master´s thesis would offer an opportunity to evaluate the project in its final stages. In addition, the framework could also be used to orchestrate new and different service experiements in the current schools participating in Buddyschool. By allowing to the users to implement the framework in a flexible way to best match their projects, the current evaluation framework would offer Migrant Youth Helsinki an organized baseline on which to build their experiments´ objectives and evaluations.

7.2 Limitations of the Research

The high design-related complexity involved in the research, time constrains, and contextual factors contributed in setting limits to this research.

As described before, the time constrains represented one of the major elements of limitation during this master's thesis. Time constraints were both present in the background research, as well as in the actual fieldwork. In fact, the topic featured in this master's thesis could be easily expanded to incorporate more accurate and exhaustive research. In addition, working with the school environment, the time that the actors in this field could dedicate to this research was restricted.

Another contextual factor that played a big role in limiting the work was the sudden outbreak of COVID-19 that started halfway through the process. Even though most of the research work was completed before the outbreak, the closing of schools and the rapid shift towards more pragmatic issues from the actors involved in the research prevented opportunities for validation. Although the impact on the actual data gathering for the research was mild, the impact that COVID-19 had on the progress of the master's thesis was more significant. The global state of emergency and the sudden switch to remote working represented a limitation, at least in the initial period of adjustement.

Remote work only allowed for validations conducted online. The framework is validated by two of the experts interviewed during the background research. Their feedback was valuable in order to further implement the proposal, but there is a lack of structured validation sessions. In addition to that, the validation process with the staff from Migrant Youth Helsinki revealed to be more difficult than expected. Although trials were made, it was hard to translate the process and the reasons behind the framework into a language that would not feature complex design terminology and references. This prevented the staff from Migrant Youth Helsinki to be able to make relevant comments on the adaptation of the framework to Buddyschool. Therefore, the current work is not thoroughly validated by external opinions or third parties' research work.

As presented throughout different chapters of this master´s thesis, there were several limitations connected to the methods implemented to gather and analyze data. For instance, the fact that part of the fieldwork´s research and one experts´ interview were conducted in languages other than English (Italian and Finnish) highlights how the data collected is subject to the interpretation of the person translating it. Furthermore, the fieldwork not always developed as planned, and quick adaptations of the methods utilized were necessary on several occasions. This might have impacted the quality of the data gathered. In general, a personal interpretation of data was a consistent factor throughout the process, which might have led to misinterpretations of fieldnotes or behaviors observed in the schools.

Chapter six, *6. Research Conclusions and Proposal*, also underlined how a wrong assumption shaped the questions asked during interviews in the school visit: after the interviews with Migrant Youth Helsinki staff, the expectation was that everybody involved in the project would have a similar level of knowledge on the co-design process underpinning Buddyschool. This proved wrong and needed to be addressed with a change in the questions during the actual

fieldwork, but it also proved to be a precious and valuable learning, both on a personal and research level.
Another limitation was having access only to selected actors in the Buddyschool process, for instance preventing the opportunity to include in this research the teachers who participated in the co-design workshops.

7.3 Personal Reflections

In this section I will be presenting some personal reflections on the process. The first-person pronoun and a more informal language are utilized here as a way of emphasizing the personal nature of the considerations expressed below.
Looking back at the process that led me here, it is inevitable to recognize the immense value that this project brought to me. First and foremost, it was interesting to see how the thesis developed and changed throughout all the process, following different directions, then merging in one path, and then splitting again. This work developed in ways that I was not expecting and had not planned to. An example is the important role that the literature review ended up playing in my work. At the beginning, I planned to refer to and base my proposal much more on the fieldwork, but in the end the background research and the literature review turned out to be the main stones on which to base my proposal.
Totally unexpected was also the situation brought about by the COVID-19 pandemic. Looking at the process now, it is insightful to see how the work was shaped and modeled by the different moments of my professional and personal life during the months of the research. Bearing with remote ways of working and the mental stress caused by the outbreak was not easy and it certainly had an impact on my working style and the progress of the thesis. Nevertheless, this also represented an opportunity for personal growth.
Considerations must be made also on the role of the designer and the role of the partner. At the beginning of the project, I sought the support of a partner because I wanted to access their resources and help during the thesis process. In addition to that, I was drawn towards the implementation of an adaptation of Action Research process and methodology. Although I am grateful for the opportunity to work with Migrant Youth Helsinki and I fully acknowledge the benefits received by working with them, as well as the perfect opportunity represented by Buddyschool as a case study for my topic, I also recognize that working so closely with a partner might have not impacted my research in the ways I was expecting to. Although we always were in close contact throughout most of the process, especially towards the end it became clear that the interests of the staff were no longer close to my personal interests in the topic. Seeking for feedback required long preparations where I would try to "translate" the framework in the easiest way possible to gather opinions and validations from the staff. Unfortunately, maybe also facilitated by the fact that no meetings in person were held in the validation phase, most of these attempts failed. It is interesting to highlight how one of the patters recognized through my thesis research can also be observed in the thesis work itself: in fact, there was a failed translation process from experiential knowledge (the one I have about the framework) to narrative knowledge (the one Migrant Youth Helsinki staff has about the framework). This failed translation might have been caused by my lack of skills in making the framework more understandable to them, as well as the high design literacy and a high familiarity with service design and co-design approaches needed in order to fully and holistically understand the items in the framework.
In general, through this thesis I discovered a great interest for this topic, that I hope I will be able to research further in the future. I am grateful for the process because it offered an opportunity for personal exploration on various fronts. Anne-Marie Willis (2006) says: "we are designed by our designing and by that which we have designed"(Willis, 2006, p. 80). I think this brief sentence carries a very powerful meaning that I

have been reflecting on very often, ever since I joined a design school. I believe that we are indeed designed by our own designing and after finishing my thesis I feel like I know a bit more about myself, my design practice, and my thinking process. Above all, I learnt and had the confirmation that design makes my life meaningful, and that this is my life choice and my career. As a young student who is now entering into the professional world, this self-discovery journey through my thesis was an important step to take in order to explore further my capabilities and skills, understanding what my strengths are and on what I still need to work on. Beyond the results of the research, these are the biggest learnings and takeaways from this six-months journey.

7.4 Conclusions

With co-design practices becoming of increasing interest for public and private sectors companies, the need to further research the practices surrounding the co-development process is growing. One of the areas that is receiving a continuously growing attention is service and co-design evaluation.

Although previous literature has focused on researching evaluation practices and co-design benefits, there is still a need for further exploring the topic (Hoyer et al., 2010; Ostrom et al., 2010; Steen et al., 2011).

Specifically, the context in which the proposal of this master's thesis is introduced is the lack of further research on co-design evaluation. This research work argues the necessity of not only giving space to outcome-focused service evaluation, but also to dedicate resources to more process-centric indicators that highlight the relation between the co-design development process and the service performance. Therefore, the design proposal chosen for this master´s thesis is an co-design evaluation framework for service development projects. When the general framework is applied to a specific case study or project, the metrics offer opportunities for evaluating the collaborative performance in the development phase and the effects of such performance on the service results in the implementation phase.

The general co-design evaluation framework clusters metrics and indicators gathered through the Buddyschool case study research and the literature review into six meaningful categories. Every area of interest represents a specific moment or item featured in the co-design process and comprehends several relevant indicators.

The framework adds to the previous research because it compiles an organized selection of metrics related and relevant to the co-design process. The framework should be seen as a first step towards an attempt to establish a more structured and detailed future research opportunities. While the current framework does not provide the tools for value judgments, it does accomplish the task of sparking considerations on the process and on its impacts on the service performance. By doing this, the framework can be seen as a tool for reflecting and making changes *in itinere* with the help of leading indicators, but also for analyzing and exploring the results after the service implementation through lagging indicators (Manuele, 2009).

Although the creation of the proposal is closely intertwined with research that encompasses a specific case study, the general framework can potentially be applied to different and diverse projects or services, bringing additional value and more occasions for further research.

In addition to exploring the opportunity for co-design evaluation, the case study research and the background research conducted in this master´s thesis produced other essential learnings.

For instance, the Buddyschool case study research, with the school visit and the interviews with Migrant Youth Helsinki staff, offered an opportunity to investigate the role of the participants and actors in the co-design process. The research highlighted how the teachers that did not participate in the co-development process (Everyday Teachers) were not aware of its value

and dynamics, resulting in their inability to connect the service results to the co-design process. On the contrary, the teachers that participated in the co-development process held experiential knowledge of it, since they participated in first person. The awareness of Teachers Co-Designers to the co-design values and opportunities might also transform them into Ambassadors, willing to advocate for the co-design process and its benefits with their Everyday Teachers colleagues.

These findings can be collocated in a wider discourse about the dissemination of learnings that can occur from actor to actor during and after the co-design process. In fact, it was observed how the Ambassadors and the Teachers Co-Designers are the ones that can foster knowledge and awareness about the co-development process in the Everyday Teachers: by doing so, the experiential knowledge of the Ambassadors and the Teachers Co-Designers is transferred to the Everyday Teachers in the form of narrative knowledge.

Finally, one of the major learnings produced through the background research, and specifically through the experts´ interviews, is that evaluation practices within the service and co-design field are complex to conduct and even to conceptualize. Paradoxically, the more an organization is using design on a higher level, the more it is extremely complicated to isolate its effects and doings. In addition to that, co-design processes are always context-dependent, and they take into consideration a wide array of different factors, making it impossible to detect any "one size fits all" metrics to utilize when aiming at evaluating collaborative performances and their effects on the service results.

Despite the complexity of the topic, this master´s thesis wants to argue for the importance of researching co-design evaluation.

In fact, finishing where this research started, Sanders & Stappers (2008) underline that "the application of participatory design practices . . . will change design and may change the world" (Sanders & Stappers, 2008, p. 9). And indeed, if the world might change, humans need to be able to understand these changes, their meanings, their impacts, and their effects. Evaluative and reflective practices play a key role in this landscape of continuous change because they allow designers and key stakeholders to reflect on the process and being accountable for it, fostering improvements and learnings, as well as helping in advocating for a bigger role for co-design.

Blank Page

References

Adil, M. (2013). SERVQUAL and SERVPERF: A Review of Measures. *Global Journal of Management and Business Research Marketing, 13*(6), 65-76.

Adzic, G., Domingues, I., & Berndtsson, J. (2014). *Getting the most out of impact mapping*. InfoQ. Retrieved January 20, 2020 from https://www.infoq.com/articles/most-impact-mapping/

Alam, I. (2002) An exploratory investigation of user involvement in new service development. *Journal of the Academy Marketing Sciences, 30*. https://doi.org/10.1177/0092070302303006

Arthur, S., Nazroo, J., & O'Connor, W. (2003). Designing Fieldwork Strategies and Materials. In *Qualitative research practice: A guide for social science students and researchers* (Vol. 41). SAGE Publications.

Banke-Thomas, A. O., Madaj, B., Charles, A., & van den Broek, N. (2015). Social Return on Investment (SROI) methodology to account for value for money of public health interventions: A systematic review. *BMC Public Health*, *15*(1), 582. https://doi.org/10.1186/s12889-015-1935-7

Björklund, T., Hannukainen, P., & Manninen, T. (2018). Measuring the impact of design, service design and design thinking in organizations on different maturity levels. In A. Meroni, A. M. Ospina Medina, & B. Villari (Eds.), *Proceedings of the ServDes2018 Conference,* 500-511

Blomkamp, E. (2018). *Sharing the principles of co-design*. Medium. Retrieved January 15, 2020 from https://medium.com/@emmablomkamp/sharing-the-principles-of-co-design-4a976bb55c48

Burton, P. (2009). Conceptual, Theoretical and Practical Issues in Measuring the Benefits of Public Participation. *Evaluation*, *15*(3), 263–284. https://doi.org/10.1177/1356389009105881

Carpini, M. X. D., Cook, F. L., & Jacobs, L. R. (2004). PUBLIC DELIBERATION, DISCURSIVE PARTICIPATION, AND CITIZEN ENGAGEMENT: A Review of the Empirical Literature. *Annual Review of Political Science*, *7*(1), 315–344. https://doi.org/10.1146/annurev.polisci.7.121003.091630

Chess, C. (2000). Evaluating Environmental Public Participation: Methodological Questions. *Journal of Environmental Planning and Management*, *43*(6), 769–784. https://doi.org/10.1080/09640560020001674

Coghlan, D., & Brannick, T. (2005). *Doing Action Research in Your Own Organization* (Second Edition). Sage Publications.

Coryn, C. L. S., Wilson, L. N., Westine, C. D., Hobson, K. A., Ozeki, S., Fiekowsky, E. L., Greenman, G. D., & Schröter, D. C. (2017). A Decade of Research on Evaluation: A Systematic Review of Research on Evaluation Published Between 2005 and 2014. *American Journal of Evaluation*, *38*(3), 329–347. https://doi.org/10.1177/1098214016688556

Costello, P. J. M. (2011). *Effective Action Research. Developing Reflective thinking and Practice* (2nd edition). Continuum International Publishing Group.

Danish Design Centre. (2001). *The Design Ladder: Four steps of design use*. Dansk Design Center. Retrieved February 3, 2020 from https://danskdesigncenter.dk/en/design-ladder-four-steps-design-use

Denscombe, M. (1998). *The Good Research Guide* (2nd Edition). Open UniversityPress, Suffolk.

Denscombe, M. (2009). *Ground Rules for Social Research: Guidelines for Good Practice*. McGraw-Hill Education.

Drew, C. (2019). *An Iterative, Experience and Practice-led Approach to Measuring Impact*. Medium. Retrieved February 11, 2020 from https://medium.com/touchpoint/an-iterative-experience-and-practice-led-approach-to-measuring-impact-db8abb7939e2

Foglieni, F., & Holmlid, S. (2017). Determining Service Value: Exploring the Link Between Value Creation and Service Evaluation. *Service Science*, *9*(1), 74–90. https://doi.org/10.1287/serv.2016.0164

Foglieni, F., Villari, B., & Maffei, S. (2018). *Designing Better Services*. Springer International Publishing. https://doi.org/10.1007/978-3-319-63179-0
Frayling, C. (1993). Research in Art and Design. *The Royal College of Art Research Papers*, *1*(1).

Gupta, S., & Vajic, M. (2000). The Contextual and Dialectical Nature of Experiences. In *New Service Development: Creating Memorable Experiences*, 33–51. SAGE Publications, Inc. https://doi.org/10.4135/9781452205564.n2

Gustavsen, B. (2008). Action research, practical challenges and the formation of theory. *Action Research*, *6*(4), 421–437. https://doi.org/10.1177/1476750308094130

Heller, C. (2017). *Designing a Way to Measure the Impact of Design*. Standford Social Innovation Review. Retrived January 15, 2020 from https://ssir.org/articles/entry/designing_a_way_to_measure_the_impact_of_design

Helsingin Kaupunki. (2019). *New tools for Helsinki to evaluate the impact of design*. Helsingin Kaupunki. Retrieved March 20, 2020 from https://www.hel.fi/uutiset/en/kaupunginkanslia/new-tools-to-evaluate-the-impact-of-design

Hickman, L. A. (1998). Dewey's Theory of Inquiry. In *Reading Dewey: Interpretations for a Postmodern Generation*, 166–186. Indiana University Press.

Hiekkavuo, A. (2017). *Population with foreign background in Helsinki 2016*. City of Helsinki Urban Facts. Retrieved February 15, 2020 from https://www.hel.fi/hel2/tietokeskus/julkaisut/pdf/17_01_16_Tilastoja_2_Hiekkavuo.pdf

Hoyer, W. D., Chandy, R., Dorotic, M., Krafft, M., & Singh, S. S. (2010). Consumer Cocreation in New Product Development. *Journal of Service Research*, *13*(3), 283–296. https://doi.org/10.1177/1094670510375604

Kanter, R. M. (1994). Collaborative Advantage: The Art of Alliances. *Harvard Business Review*, *July–August*.

Kleinsmann, M. (2006). *Understanding collaborative design*. Doctoral Thesis. Technische Universiteit Delft. Retrieved February 16, 2020 from https://repository.tudelft.nl/islandora/object/uuid%3A0a7a57d4-c846-4458-a59f-24c25acbafa9

Kristensson, P., Magnusson, P. R., & Matthing, J. (2002). Users as a Hidden Resource for Creativity: Findings from an Experimental Study on User Involvement. *Creativity and Innovation Management*, *11*(1), 55–61. https://doi.org/10.1111/1467-8691.00236

Kujala, S. (2003). User involvement: A review of the benefits and challenges. *Behaviour & Information Technology*, *22*(1), 1–16. https://doi.org/10.1080/01449290301782

Legard, R., Keegan, J., & Ward, K. (2003). In-depth Interviews. In *Qualitative research practice: A guide for social science students and researchers* (Vol. 41). SAGE Publications. http://choicereviews.org/review/10.5860/CHOICE.41-1319

Manuele, F. A. (2009). Leading and Lagging Indicators: Do they add value to the practice of safety? *Professional Safety, December*.

Markkanen, K. (2017). *Immigrant youth are designing services for themselves together with the city: Migrant Youth Helsinki*. Design Stories from Helsinki. Retrieved December 15, 2019 from https://www.muotoilutarinat.fi/en/project/migrant-youth-helsinki/
Minassians, H. P. (2015). Network governance and performance measures: Challenges in collaborative design of hybridized environments. *International Review of Public Administration*, *20*(4), 335–352. https://doi.org/10.1080/12294659.2015.1088689

Moultrie, J., & Livesey, F. (2009). International design scoreboard: initial indicators of international design capabilities. IFM Management Technology Policy, University of Cambridge. Design Council. United Kingdom.

NSW Council of Social Service. (2017). *Principles of Co-design*. Retrieved July 25, 2020 from https://www.ncoss.org.au/sector-hub/sector-resources/principles-of-co-design/

Ostrom, A. L., Bitner, M. J., Brown, S. W., Burkhard, K. A., Goul, M., Smith-Daniels, V., Demirkan, H., & Rabinovich, E. (2010). Moving Forward and Making a Difference: Research Priorities for the Science of Service. *Journal of Service Research*, *13*(1), 4–36. https://doi.org/10.1177/1094670509357611

Ostrom, A. L., Parasuraman, A., Bowen, D. E., Patrício, L., & Voss, C. A. (2015). Service Research Priorities in a Rapidly Changing Context. *Journal of Service Research*, *18*(2), 127–159. https://doi.org/10.1177/1094670515576315

Papas, N., O'Keefe, R. M., & Seltsikas, P. (2012). The action research *vs* design science debate: Reflections from an intervention in eGovernment. *European Journal of Information Systems*, *21*(2), 147–159. https://doi.org/10.1057/ejis.2011.50

Parasuraman, A. P., Zeithaml, V. A., & Berry, L. L. (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing*, 11.
Pawson, R., & Tilley, N. (1997). *Evaluation for the 21st Century: A Handbook*. SAGE Publications, Inc. https://doi.org/10.4135/9781483348896

Pirinen, A. (2016). The Barriers and Enablers of Co-design for Services. *International Journal of Design, 10*(3), 27-42.

Prahalad, C. K., & Ramaswamy, V. (2004). Co-creation experiences: The next practice in value creation. *Journal of Interactive Marketing*, *18*(3), 5–14. https://doi.org/10.1002/dir.20015

Reason, P., & Bradbury, H. (2006). Introduction: Inquiry and Participation in Search of a World Worthy of Human Aspiration. In *Handbook of Action Research. The Concise Paperback Edition*. SAGE Publications.

Ritchie, J. (2003). The Applications of Qualitative Methods to Social Research. In *Qualitative research practice: A guide for social science students and researchers*, *41*. SAGE Publications.

Sanders, E. (2002). From user-centered to participatory design approaches. In J. Frascara (Ed.), *Design and the Social Sciences*, 1–8.

Sanders, E. B.-N., & Stappers, P. J. (2008b). Co-creation and the new landscapes of design. *CoDesign*, *4*(1), 5–18. https://doi.org/10.1080/15710880701875068

Scriven, M. (1991). *Evaluation Thesaurus* (Fourth Edition). SAGE Publications.

Steen, M. (2013). Co-Design as a Process of Joint Inquiry and Imagination. *Design Issues*, *29*(2), 16–28. https://doi.org/10.1162/DESI_a_00207

Steen, M., Manschot, M., & Koning, N. D. (2011). Benefits of Co-design in Service Design Projects. *International Journal of Design*, *5*(2), 53-60.

Suominen, S., & Pöyry-Lassila, P. (2013). Effectiveness of Co-Design Intervention – Adopting Service Co-Development Thinking. In Riitta Smeds & Olivier Irrmann (eds.), *Proceedings of the Boundary-Crossing Conference on Co-Design in Innovation, Espoo*.

Trischler, J., Pervan, S. J., Kelly, S. J., & Scott, D. R. (2018). The Value of Codesign: The Effect of Customer Involvement in Service Design Teams. *Journal of Service Research*, *21*(1), 75–100. https://doi.org/10.1177/1094670517714060

Usertesting. (2018). *Customer Experience Metrics: A Brief Guide on How to Measure CX*. Usertesting Blog. Retrieved January 20, 2020 from https://www.usertesting.com/blog/customer-experience-metrics/

Vargo, S. L., & Lusch, R. F. (2004). Evolving to a New Dominant Logic for Marketing. *Journal of Marketing*, *68*(1), 1–17. https://doi.org/10.1509/jmkg.68.1.1.24036

Visser, F. S., Stappers, P. J., van der Lugt, R., & Sanders, E. B.-N. (2005). Contextmapping: Experiences from practice. *CoDesign*, *1*(2), 119–149. https://doi.org/10.1080/15710880500135987

Voss, C. A. (1992). Measurement of Innovation and Design Performance in Services. *Design Management Journal (Former Series)*, *3*(1), 40–46. https://doi.org/10.1111/j.1948-7169.1992.tb00586.x

Willis, A.-M. (2006). Ontological Designing—Laying the Ground. *Design Philosophy Papers*, *4*(2), 69-92. https://doi.org/10.2752/144871306X13966268131514

Yin, Y., Qin, S., & Holland, R. (2011). Development of a design performance measurement matrix for improving collaborative design during a design process. *International Journal of Productivity and Performance Management*, *60*(2), 152–184. https://doi.org/10.1108/17410401111101485

Yin, Y., Shengfeng, Q., & Holland, R. (2009). *Using Design Performance Measurement as a Strategy to Improve Collaborative Design Performance*.

List of Figures & Tables

Figures

Tables

Blank Page

Appendix

Appendix 1. Buddyschool survey for teachers.

Vaikuttavuus

Wepropol kysely, vastauksia 40.

-Kysymykset laadittu yhdessä opettajien kanssa

- Kyselyssä kysyttiin vertaisohjaajien ja ohjattavien oppimisen, käyttäytymisen ja vuorovaikutuksen parantumiseen liittyviä väittämiä. Lisäksi kysymyksiä liittyen kouluyhteisöön. Opettajat vastasivat väittämiin liukumalla 0-5.
O= ei yhtään, 5= todella paljon

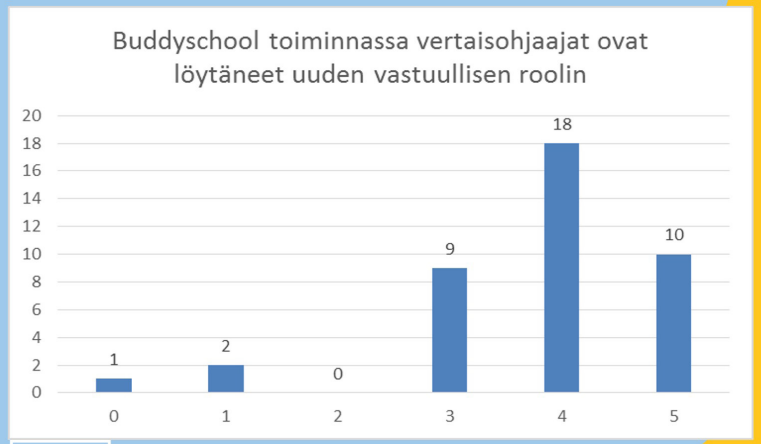
-Yhteenveto vertaisohjaajien vastauksista-Buddyschoolissa nuoret saavat vastuullisen roolin, joka kehittää vuorovaikutustaitoja ja parantaa koulumotivaatiota**. > oppimistulosten parantuminen**

-Buddyschool lisää oppilaiden ja opettajien vuorovaikutusta**> hyvinvoiva kouluyhteisö**

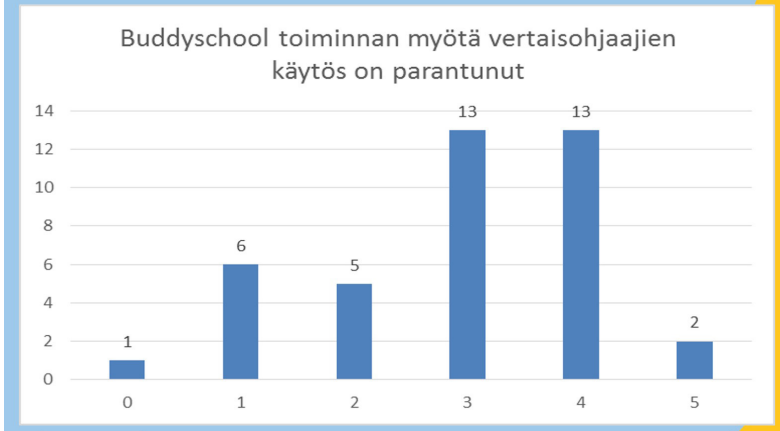
Helsinki

OHJAAJAT

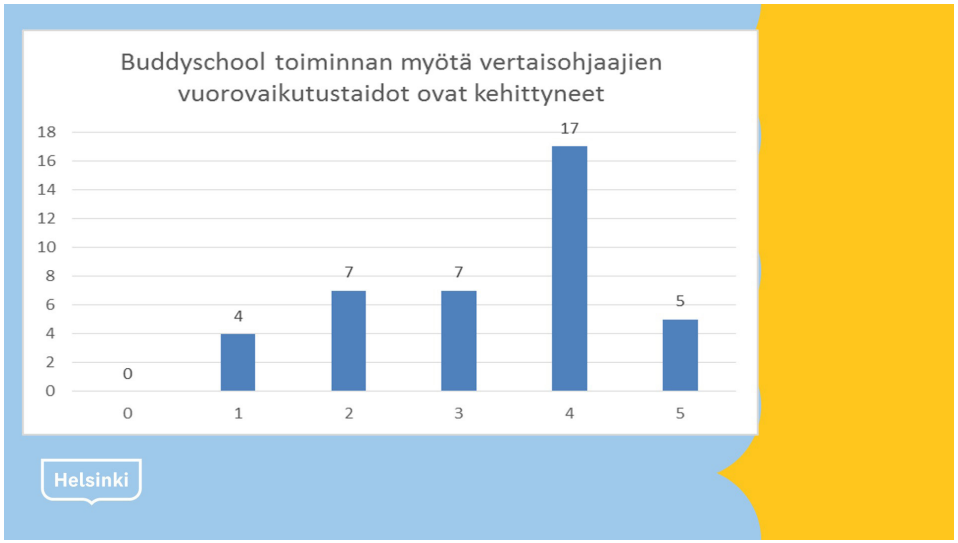
Helsinki

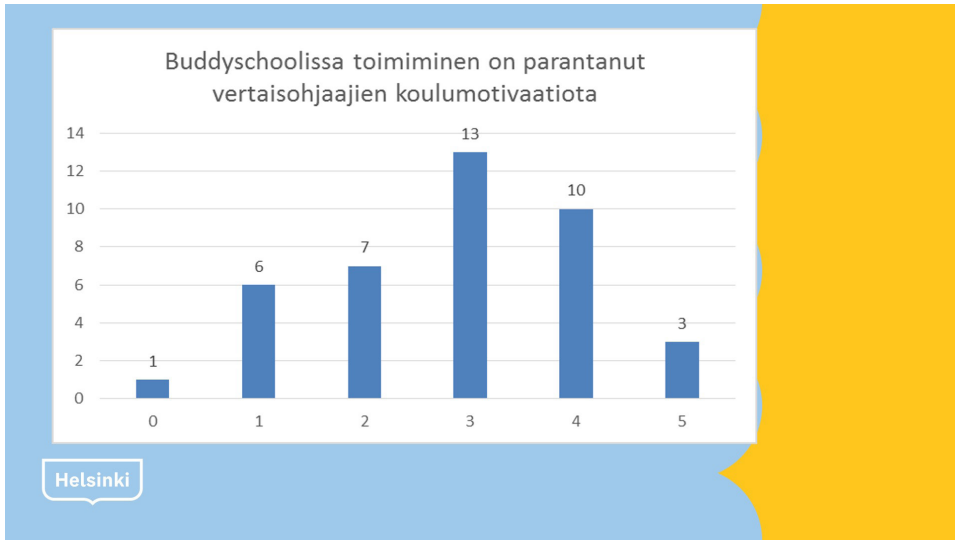


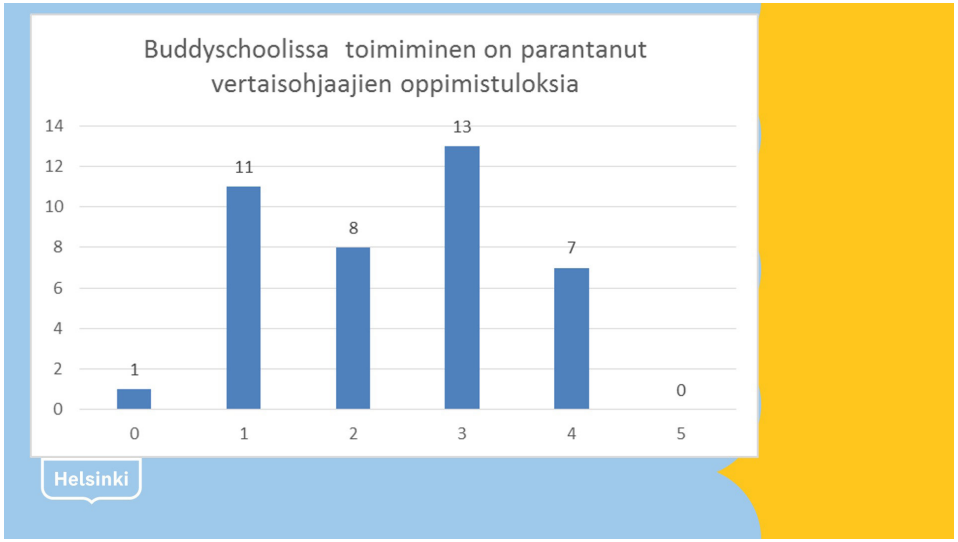
Helsinki



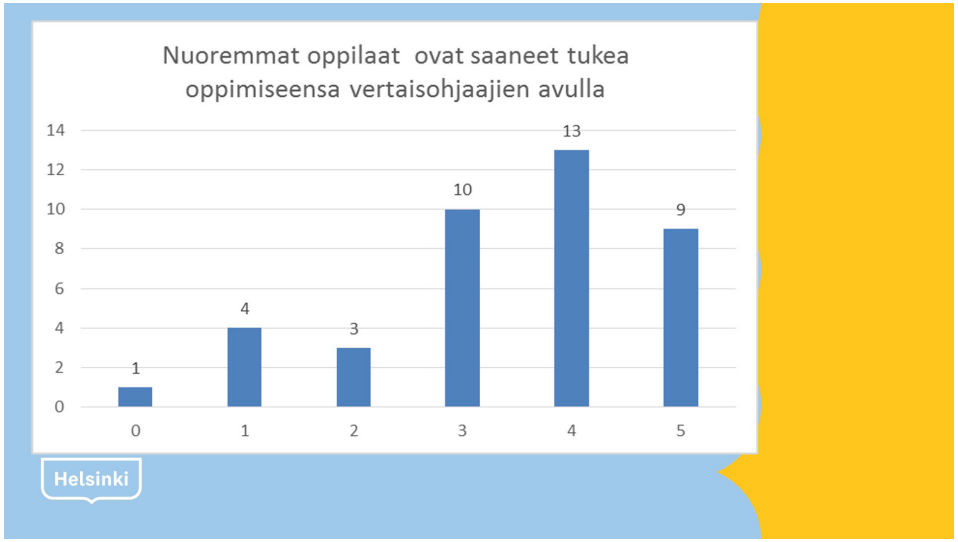
Helsinki

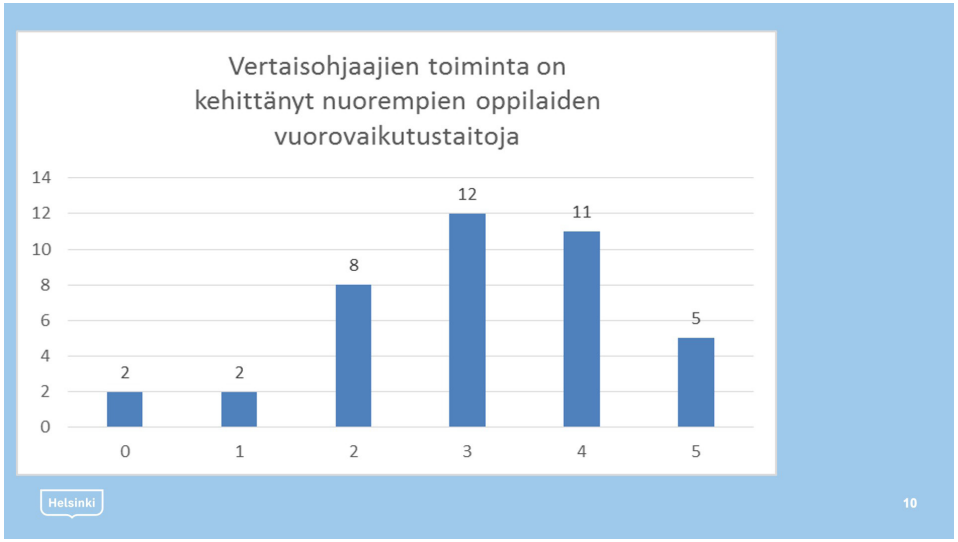

Buddyschool toiminnan myötä vertaisohjaajien vuorovaikutustaidot ovat kehittyneet
18
16
14
12
10
8
6
4
2
0
0
4
7
7
17
5
0
1
2
3
4
5
Helsinki


Buddyschoolissa toimiminen on parantanut vertaisohjaajien koulumotivaatiota
14
12
10
8
6
4
2
0
1
6
7
13
10
3
0
1
2
3
4
5
Helsinki

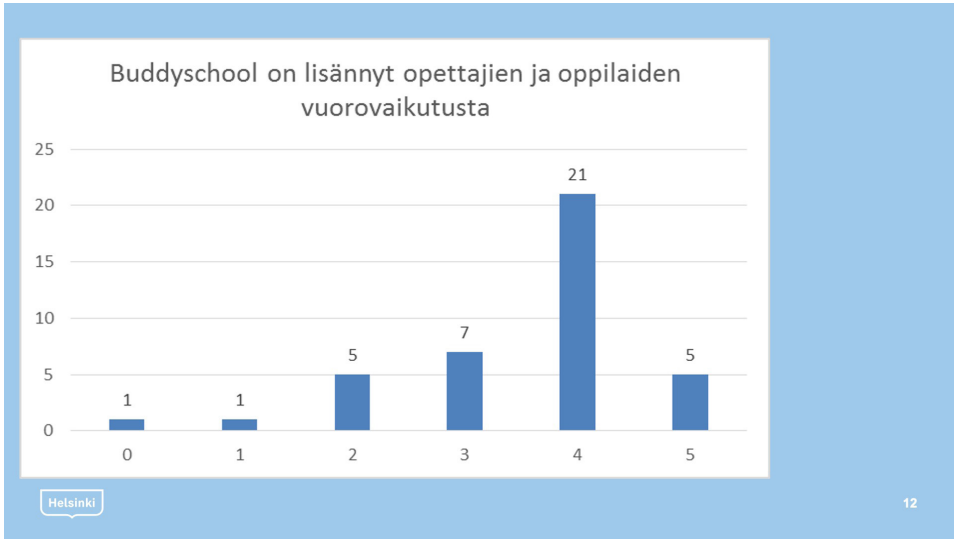

Buddyschoolissa toimiminen on parantanut vertaisohjaajien oppimistuloksia
14
12
10
8
6
4
2
0
1
11
8
13
7
0
0
1
2
3
4
5
Helsinki

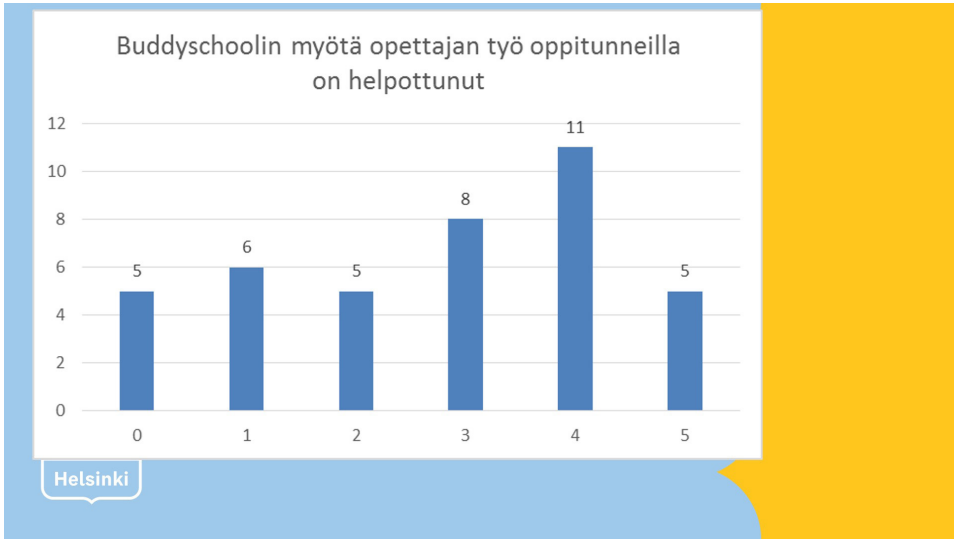

OHJATTAVAT
Helsinki


Nuoremmat oppilaat ovat saaneet tukea oppimiseensa vertaisohjaajien avulla
14
12
10
8
6
4
2
0
1
4
3
10
13
9
0
1
2
3
4
5
Helsinki


Vertaisohjaajien toiminta on kehittänyt nuorempien oppilaiden vuorovaikutustaitoja
14
12
10
8
6
4
2
0
2
2
8
12
11
5
0
1
2
3
4
5
Helsinki
10


KOULUYHTEISÖ
Helsinki


Buddyschool on lisännyt opettajien ja oppilaiden vuorovaikutusta
25
20
15
10
5
0
1
1
5
7
21
5
0
1
2
3
4
5
Helsinki
12



Opettajien arviointia:

”Buddyschool on toimiva menetelmä/vaihtoehto silloin, kun omassa luokassa oppiminen on jostain syystä haasteellista/vaikeaa. Se antaa mahdollisuuden nähdä oppilas uudessa, myönteisessä valossa, varsinkin jos palaute usein on ollut negatiivista”

”Niille, joille koulu ja työskentely on jo valmiiksi helppoa, buddyschool ei ehkä aina edes sovi tai ei tunnu mukavalta”

“Mukana oleville tärkeä tehtävä, joka näkyy itseluottamuksen kasvuna”

”BS tuo kouluun lisää viihtyvyyttä, sitoutumista ja oppimisen iloa!”

Helsinki

14

Blank Page

Blank Page